



The impact of clinical supervision on
counsellors and therapists, their practice
and their clients: a systematic review of
the literature

bacp

British Association for
Counselling & Psychotherapy

Sue Wheeler & Kaye Richards

Contents

1.	Executive Summary	Page 3
2.	Background	Page 4
2.1	Supervision: an introduction	4
2.2	Supervision research reviews	4
2.3	A definition of supervision	4
2.4	A history of supervision	4
2.5	Models of supervision and the role of the supervisor	4
2.6	Supervision research	5
2.7	Answering questions through systematic review techniques	5
2.8	Defining the review question	5
2.9	Related supervision research reviews	5
3.	Methodology	Page 7
3.1	Structure of the review tasks	7
3.2	Phase One: scoping and retrieval of relevant studies	9
3.2.1	Scope of the review: defining inclusion and exclusion criteria	9
3.2.2	Search strategy	9
3.2.3	EPPI-Reviewer	9
3.2.4	Reviewing, recording and rating relevance of methodology of each study	10
4.	Results	Page 11
4.1	An overview of the titles of studies included and excluded	11
4.1.1	An overview of titles of included studies	11
4.1.2	An overview of titles of studies unobtainable and excluded	11
4.2	Supervision impact categories	12
4.3	Methodological approaches and measures used in the studies	12
4.4	The supervisees, supervisors and clients in the studies	12
4.4.1	The supervisees	12
4.4.2	The supervisors	13
4.4.3	The clients	13
4.5	Quality of the evidence	13
4.5.1	Quality of the evidence and study limitations	13
4.5.2	Examples of methodological problems	15
4.6	Research evidence	17
4.6.1	Self-awareness	17
4.6.1a	An overview	17
4.6.1b	The evidence	17
4.6.1c	A summary of the evidence	29
4.6.2	Skill development	29
4.6.2a	An overview	29
4.6.2b	The evidence	29
4.6.2c	A summary of the evidence	31

4.6.3	Self-efficacy	31
	4.6.3a An overview	31
	4.6.3b The evidence	31
	4.6.3c A summary of the evidence	32
4.6.4	Transfer of understanding from supervision to therapy	32
	4.6.4a An overview	32
	4.6.4b The evidence	32
	4.6.4c A summary of the evidence	33
4.6.5	Client outcome	34
	4.6.5a An overview	34
	4.6.5b The evidence	34
	4.6.5c A summary of the evidence	34
5.	Conclusions and implications for research and practice	Page 35
6.	Research Articles	Page 37
6.1	Abstracts of articles/studies included in the review	37
6.2	Abstracts of articles/studies unable to obtain in Phase Two	62
6.3	Overviews of relevant supervision research reviews	66
7.	References	Page 69
7.1	Articles included in the review	69
7.2	Related supervision research reviews	69
7.3	Articles include in Phase Two but unobtainable	69
7.4	Articles excluded during Phase Two	70
7.5	Articles from Phase One Screening	71
8.	All References	Page 72
9.	Appendix	Page 75
9.1	Data-extraction form	75

List of tables and figures

Figure 1:	Flowchart of Phase One and Phase Two: an overview of literature search/retrieval, initial screening of studies and data extraction	8
Table 1:	Exclusion criteria for the screening of articles/studies	10
Table 2:	A list of the 25 articles/studies included in this review	11
Table 3:	The frequency of supervision in initial-impact categories	12
Table 4:	An overview of supervision impacts for each study	13
Table 5:	An overview of general methodological approaches to studies	14
Table 6:	An overview of types of methods used to collect data	15
Table 7:	Instruments used in the reviewed studies	16
Table 8:	Research evidence for self-awareness	18
Table 9:	Research evidence for skill development	19
Table 10:	Research evidence for self-efficacy	22
Table 11:	Research evidence for transfer of understanding from supervision to therapy	26
Table 12:	Research evidence for client outcome	28

Section 1: Executive summary

Overview

The British Association for Counselling and Psychotherapy (BACP) commissioned a systematic review of the research evidence on the impact of supervision on counsellors and psychotherapists, their practice and their clients. The objectives of this review were to systematically locate, appraise and synthesise evidence from research studies in order to obtain a reliable overview of the impact of clinical supervision. Detailed inclusion and exclusion criteria were agreed. EPPI-Reviewer software was used to organise and manage the analysis of the articles that met the inclusion criteria. This paper reports on the findings of this review, from articles published in this area since 1980, and reviews 25 individual published studies. The quality of evidence is variable, but supervision is consistently demonstrated to have some positive impacts on the supervisee.

Methods

The review identified and summarised the evidence relating to review questions. The inclusion criteria for the review stipulated that the research had to examine some level of impact and not just be about the supervisee's satisfaction with supervision. After searches of electronic databases, hand-searches of relevant books and journals, internet searches and citation tracking, more than 8,000 studies were screened and 33 studies were deemed appropriate for inclusion in the review. Eight studies were unobtainable at the time; therefore 25 are included in the final review. A critical appraisal of each paper by up to three independent reviewers identified the relevant results from each study. The studies included a range of methodological approaches – quantitative, qualitative and mixed methods. The results from these studies are presented in summary tables, along with a narrative report of them and consideration of methodological limitations.

The findings of impacts of supervision

The systematic review found:

- limited evidence that supervision can enhance the self-efficacy of the supervisee (seven quantitative studies);
- limited evidence that supervision has a beneficial effect on the supervisees, the client and the outcome of therapy (one RCT);
- limited evidence that supervision that focuses on the working alliance can influence client perception of this and enhance treatment outcome in the brief psychotherapeutic treatment of depression (one RCT);
- limited evidence that clients treated by supervised therapists are more satisfied than those treated by unsupervised therapists (one RCT);
- limited evidence that both skill and process supervision have the same positive impact on client outcome (one RCT);
- limited evidence that counselling and psychotherapeutic skills develop through supervision (seven quantitative studies, one mixed-method study, two qualitative studies);

- preliminary evidence that supervisee self-awareness increases as a result of counselling and psychotherapy training, and that some of that development may be attributed to supervision (one qualitative study, one cross-sectional survey and one longitudinal study);
- tentative evidence to support the assertion that learning in supervision is transferred to practice (one quantitative study, three mixed-method studies and one single case qualitative study);
- tentative evidence that there is thematic transfer of an appropriate kind from supervision to therapy (one single case qualitative study);
- tentative evidence to suggest that trustworthiness of the supervisor is an important factor in effective supervision (one quantitative study);
- tentative evidence to suggest that supervisees perceived individual supervision as safer than group supervision in promoting their personal growth (one qualitative study);
- tentative evidence to suggest that the timing of supervision can influence what is dealt with (one mixed method study).

The findings of methodological issues

The systematic review found that:

- most of the evidence is related to the impact of supervision for trainees – there is a clear need to examine the impact of supervision on qualified and experienced practitioners;
- most of the evidence is about the impact of supervision on the supervisee – there is a clear need for more research about the impact of supervision on client outcome;
- there is a need to examine the longer term impact of supervision with more robust longitudinal studies;
- many studies reviewed have a range of methodological weaknesses – there is a need for the development of robust methodology for supervision research;
- there is very little research on supervision in the UK, thus a strategic UK supervision research agenda is urgently required.

Section 2: Background

1.1 Supervision: an introduction

Supervision of counselling and psychotherapy practice is widely promoted as an essential aspect of ethical and effective therapy, and is seen as the cornerstone of continuing professional development. The Ethical Framework for Good Practice of Counselling and Psychotherapy (BACP, 2002) stresses the importance of supervision in supporting the practitioner to adhere to the ethical framework and in being a crucial aspect of the infrastructure that underpins professional practice. In the UK, the British Association for Counselling and Psychotherapy (BACP) requires all accredited therapists to have supervision throughout their career, and other organisations representing counsellors and psychotherapists strongly recommend supervision. Requirements for the supervision of therapists vary in different countries; in the USA, where most research into supervision is conducted, it is only trainees who routinely participate in supervision. A previous scoping search of supervision (Wheeler, 2003: see pg 6 for a brief overview) found that out of 388 studies, only 11 had been conducted in Britain and only six related to experienced practitioners, highlighting the way in which supervision research is skewed towards work abroad and to work with trainees.

2.2 Supervision research reviews

There have been several reviews of supervision literature in recent years (Kilminster & Jolly, 2000; Lambert & Ogles, 1997; Milne & James, 2000; Freitas, 2002). The general conclusion is that various aspects of supervision, the supervisor or the relationship with the supervisee have an effect on the supervisee and their understanding of the process of therapy and practice with clients. Inspection of the inclusion criteria for these reviews reveals considerable variety. For example, the Kilminster and Jolly (2000) review focused entirely on supervision in medical settings. The Lambert and Ogles (1997) comprehensive review made methodological rigour its priority, and the Milne and James (2000) review focused specifically on cognitive behavioural supervision in the health professions, with a majority of their identified studies coming from the field of learning disabilities.

2.3 A definition of supervision

There are many definitions of supervision that reflect the way in which it is understood by different professions in different contexts and countries. Supervision is a formal relationship in which there is a contractual agreement that the therapist will present their work with clients in an open and honest way that enables the supervisor to have insight into the way in which the work is being conducted. The supervisor is understood to be accountable to the professional body to which the supervisee has allegiance (Wheeler, 2003). Inskipp and Proctor (2001a: 1) describe supervision as:

“A working alliance between the supervisor and counsellor in which the counsellor can offer an account or recording of her work; reflect on it; receive feedback and, where appropriate, guidance. The object of this alliance is to enable the counsellor to gain in ethical competence, confidence, compassion and creativity in order to give her best possible service to the client.”

A classic definition of supervision that is often quoted is provided by Bernard & Goodyear (1992: 2004), who say that

“supervision is an intervention provided by a more senior member of a profession to a more junior member of that same profession. This relationship is evaluated, extends over time and has the simultaneous purposes of enhancing the professional functioning of the more junior person, monitoring the quality of professional services offered to the clients seen and serving as a gatekeeper for those who are to enter the particular profession” (2004: 8).

This description might fit current supervision practice in the USA but it does not capture the essence of the complexity of the supervision process, tasks, roles and functions in all settings. Milne (*in press*) has conducted an empirical study into the definition of supervision that has resulted in the following definition:

“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 colleague/s. The main methods that supervisors use are corrective feedback on the supervisee’s performance, teaching and collaborative goal setting. It therefore differs from related activities, such as mentoring and coaching, by incorporating an evaluative component. Supervision’s objectives are ‘normative’, ‘restorative’ and ‘formative’. These objectives could be measured by current instruments.”

2.4 A history of supervision

The supervision of psychotherapeutic practice has been documented almost since the inception of psychoanalysis, and has been integral to the training of analysts and social workers and, more recently, therapists and counsellors (Jacobs, David & Meyer, 1995). Freud supervised other analysts often through letters and conversations, unlike the formal arrangements for supervision that exist in professional organisations and training agencies today. In the early years of psychoanalytic training, trainees discussed their cases with their own analyst, in sharp contrast to the ethical prohibition on such dual relationships now.

Kadushin (1968) and Mattinson (1977) developed their own ideas about supervision, within the context of social work, which gradually began to influence other professional groups. In 1987, Hess proposed that supervisors should be trained to do that specific job, articulating ways in which the practice of therapy and supervision are different and require different skills. Good therapists are not necessarily good supervisors. During the past three decades, provision of supervision has become widespread since Ekstein and Wallerstein (1972) endorsed the practice of supervision as being essential for maintaining clinical standards.

2.5 Models of supervision and the role of the supervisor

Many models of supervision have been derived by American researchers such as Stoltenberg (1981) and Loganbill, and Hardy and Delworth (1982), and it is only in recent years that British texts have appeared to provide models of supervision for British audiences. In 1989, Hawkins and Shohet published their classic text *Supervision in the Helping Professions*, which offered a process model of supervision. Page and Wosket (2001) followed on with their cyclical model. Carroll (1996) offered an alternative model of supervision based on his doctoral research. Meanwhile, Inskipp and Proctor (2001a; b) have been writing about, and teaching, supervision for

two decades using a model both informed by Hawkins and Shohet (1989) and informing the second edition of their book.

Models of supervision provide a structure for understanding the roles, relationship, responsibilities and processes integral to the practice of supervision. For example, in the process model of supervision (Hawkins & Shohet, 2000), focus on the counsellor has equal importance to all other aspects of supervision. This is not about therapy for the counsellor, but a focus on ways in which the counsellor impacts on the client, the therapeutic relationship, the supervisor and the supervisory relationship.

The developmental model (Stoltenberg & Delworth, 1987) suggests that supervisees have different needs depending on their stage of development. Trainees are postulated to need more help, support and guidance than more experienced practitioners, and various studies have sought to confirm this supposition. Holloway (1995) describes a systems model of supervision, highlighting the many roles and responsibilities, tasks and functions of the supervisor.

2.6 Supervision research

Much research has been conducted to explore aspects of supervision in an attempt to establish an evidence base to support its continued practice. The focus of such research has included numerous perspectives such as the roles, tasks and function of supervision in different therapeutic environments and contexts, and the nature and impact of the supervisory relationship. Following Searles' (1955) first description of the parallel process, an unconscious process through which the supervisor gains insight into the therapist's relationship with the client as it is mirrored in the supervisory relationship, the process of supervision has been of considerable interest. Studies have explored the experience of supervisees and specific events in supervision, as well as supervision of particular client groups. Gender and cross-cultural communication have motivated some researchers, while supervisor characteristics, behaviours and qualities have interested others. Some research has addressed the difference between supervision of trainees and experienced practitioners and, most important of all, the effect of supervision on the process and outcome of therapy with clients. When considering a systematic review in the area of supervision, the choice of potential topics to review is plentiful.

2.7 Answering questions through systematic review techniques

Systematic reviews aim to find as much as possible of the research relevant to particular research questions, and to synthesise those findings in a form that is easily accessible to those who make policy or practice decisions. Explicit methods are used to identify what can reliably be said on the basis of these studies. In this way, systematic reviews reduce the bias that is potentially an element in other approaches to reviewing research evidence (EPPI, 2006a).

Well-formulated questions are a crucial aspect of systematic reviews, as they determine which key components to focus on in the initial search for relevant studies.

Formulating a question in terms of the types of participants, interventions, outcomes and study designs of interest will lead naturally to specifying the criteria that will be used to select studies. It is not possible to formulate an answerable question for a review without knowing some of the studies relevant to the question, and it may become clear that the questions the review addresses need to be modified in light of evidence

accumulated in the process of conducting the review. (Cochrane Collaboration, 2006).

2.8 Defining the review question

Wheeler's (2003) scoping search revealed studies on models of supervision, the relationship, experiences of supervision, good and bad supervisors, training supervisors, cross-cultural dynamics, the process, roles and many other topics. Supervisees reported how valuable supervision is, but evidence to support the assumption that it makes a difference to clinical practice and therapist wellbeing was not immediately obvious in many studies. There are pertinent studies, such as Leach, Stoltenberg, McNeill and Eichenfield (1997), and Cashwell and Dooley (2001), that focus on counsellor development and self-efficacy, but it is necessary to look closely at studies that might seem relevant to determine whether they really address any aspect of the impact on practice. Client outcome studies give an indication of the impact of supervision, but they are few.

Practitioners are well aware of the increasing demand for evidence-based practice, and service provision is increasingly dependent on clinical governance guidelines. Supervision is continually said to be a professional requirement but easy access to evidence of its efficacy is not obviously in the public domain. Having identified the gaps in the previous supervision review and recognising the need for tangible evidence of supervision efficacy, the review question was formulated as:

What impact does clinical supervision have on the counsellor or therapist, their practice and their clients?

2.9 Related supervision research reviews

This is by no means the first supervision review to be undertaken. Bambling, King, Raue, Schweitzewr and Lambert (2006) make reference to more than 40 reviews of supervision literature, but here the word 'review' can be interpreted in many ways. During the course of searching for literature relevant to the impact of supervision on the supervisees, many reviews were located, but only five appeared to be useful to inform this review.

The five main supervision research reviews that were identified in the screening process that were in part linked to the question of this review were as follows: Ellis, Ladany, Krengel & Schult, 1996; Ellis & Ladany, 1997; Freitas, 2002; Milne & James 2000; and Wheeler, 2003. In addition, a relevant meta-analysis by Whitaker (2004) was also identified. All of the review references were included in the screening, Phase One process of this review. A small number of the reviewed studies have been included here. The five reviews had diverse inclusion criteria that were not consistent with the current review. The conclusions were similar across the reviews – particularly that there are few methodologically sound studies, and thus only limited reliable claims can be made other than that practitioners both like and are satisfied with supervision. A more detailed overview of these reviews is provided below.

Ellis, Ladany, Krengel & Schult (1996) aimed to assess the status of clinical supervision research via a thorough documentation of the major methodological validity threats and to determine the extent to which supervision researchers have replicated previous studies. It investigated empirical studies in clinical supervision published from 1981 through to 1993 to assess scientific rigor and test whether the quality of methodology had improved since the review by Russell, Crimmings and Lent (1984). The inclusion criteria for the review were supervision of individual counselling, group supervision of individual counselling, case studies and peer

supervision. It included school, rehabilitation, community, mental health counselling, counselling and clinical psychology, social work, psychiatry and psychiatric nursing settings. It excluded pre-practicum supervision, teacher supervision, speech therapy and supervision of group therapy. A total number of 144 studies were reviewed (130 articles, some reporting on more than one study, and one book).

The main focus of the review was methodology and methodological limitations and errors. A few studies met the inclusion criteria for this systematic review and are discussed later. The main conclusion was that more methodological rigour is needed in supervision research. It recommends a shift to realistic field studies, close attention to avoiding unchecked Type I and Type II error rates (choice of significance level: Type I when significance level is too lax and Type II when it is too rigorous, sometimes as a result of poor sampling or design faults), medium Effect Sizes and attention to hypothesis validity.

Ellis and Ladany's (1997) review examined inferences concerning supervisees and clients in clinical supervision, which were defined as rudimentary core premises and tentative conclusions about supervision, supervisees or clients. It aimed to address the limitations of the review by Ellis et al. (1996) by a) conducting a more focused review of supervision research that incorporated a systematic evaluation of the scientific rigour of each study, b) a reinterpretation of the findings and c) the organisation of a review according to inferences. The review included studies that addressed supervision of individual counselling or therapy and empirically based articles in journals from 1981 to 1997, and focused on inferences about supervisee or client variables. It comprised case studies, school, rehabilitation, community, mental health counselling, counselling and clinical psychology, social work, psychiatry and psychiatric nursing peer supervision, but excluded pre-practicum supervision, teacher supervision, speech therapy and supervision of group therapy. A total of 104 studies were included (95 research articles, some reporting on more than one study, and one book).

Inclusion criteria were much broader than for this review. However, two of the six areas of inference overlapped with our review: inferences regarding supervisee development and inferences regarding client outcome. The focus was an intense and thorough methodological critique and some of the studies are included in this systematic review. The main conclusion was that the majority of studies were methodologically flawed, that a lot can be learned from the mistakes of others and that there are some promising instruments that might inform future research. It highlighted the gaps in what is reliably known about supervision.

Milne & James (2000) examined the impact of supervision and consultancy, and the change process occurring between participants. Twenty-eight empirical studies of the change processes occurring between at least two members of the educational pyramid (consultant, supervisor, supervisee and patient) were included. The inclusion criteria required that supervision was analysed under field conditions and that there were objective measures of learning outcomes. The vast majority of the studies refer to learning disability, children and families, and psychiatric rehabilitation and the focus is not on counselling. All the studies (three) that refer to adult psychotherapy are either included in our review or have been screened. The main conclusion of the review was that methodologies were weak and the bias in the sample towards learning disabilities makes generalisation inappropriate. There is little information about supervisors in the studies. A check-list of features of a well-designed study is included in the review.

Freitas (2002) reviewed studies between 1981 and 1997 that examined whether clinical supervision improves client outcome. Ten studies were reviewed in detail. Caveats and criticisms, as well as suggested directions for future research, were discussed. The inclusion criterion for the review was much broader than this systematic review and not all of the studies focused on clinical counselling. Some of the studies are included in this review. The review provides a methodological critique of each study and methodological weaknesses are noted. The main conclusion is that methodological problems lead to limited conclusions.

Wheeler (2003) performed a scoping search of supervision research published in the English language between 1980 and 2002. This provided a detailed descriptive map of the supervision literature. No research articles were purposely excluded from the report if some aspect of supervision was empirically investigated. Supervision related to a broad range of professions performing counselling and psychotherapy was included. Books and discursive articles that did not report on empirical research and management supervision were excluded. There were 388 studies identified but no attempt was made to analyse them in depth. Only brief details of methodology and sample size are given. The review provides references organised in categories:

- B The experience of supervision
- C Supervisory relationship
- D Events in supervision
- E Supervision process
- H Ethical considerations
- J Supervision of trainees
- K Supervision of experienced practitioners
- L Training of supervisors
- M Supervision mode
- N Cross-cultural issues in supervision
- O Effectiveness of supervision
- R Roles, tasks and function
- T Supervision in other professions
- V Gender issues in supervision
- Y Characteristics of the supervisor

The inclusion criteria for the scoping review was much broader than this review, with much looser exclusion criteria. However, studies from the scoping review that meet the inclusion criteria for the current review have been screened. The main conclusion of the scoping review was that there are many opportunities for future research into supervision that take account of the methodological difficulties encountered in previous work. It provides a convenient reference document to identify existing studies with the potential to enable future researchers to pinpoint relevant research questions (a free copy of the review is available as a PDF from BACP's Research Department).

Finally, in a meta-analysis of 10 studies, accompanied with interviews with supervisees undertaken by the author, **Whittaker (2004)** examined the extent to which clinical supervision affects supervisees' anxiety and self-efficacy. It is part of a doctoral thesis that also included interviews to test the conclusion of the meta-analysis.

Whilst at first sight this review/meta-analysis seemed to be very relevant to this systematic review, on closer scrutiny this proved not necessarily to be the case. Of the 10 studies included in the meta-analysis, four were published in refereed journals and six were doctoral dissertations. The four journal articles have been reviewed for the systematic review but it has not been possible to obtain the six dissertations. The four articles are:

Daniels, J. A. & Larson, L.M. (2001) The impact of performance feedback on counseling self-efficacy and counsellor anxiety. *Counselor Education and Supervision*, 41:120–131.

Ellis, M. V., Kregel, M. & Beck, M. (2002) Testing self-focused attention theory in clinical supervision Effects on supervisee anxiety and performance. *Journal of Counseling Psychology*, 49: 101–106.

Johnson, M. E. (1989) Effects of self-observation and self-as-a-model on counsellor trainees' anxiety and self-evaluations. *The Clinical Supervisor*, 7: 59–70.

Williams, E., Judge, A. B., Hill, C. E. & Hoffman, M. A. (1997) Experiences of novice therapists in prepracticum: trainees', clients' and supervisors' perceptions of therapists' personal reactions and management strategies. *Journal of Counseling Psychology*, 44: 390–399.

However, none of these articles have been included in this systematic review because they do not meet the inclusion criteria. In each, case studies had an experimental design that used role-play clients, simulated client material or volunteer clients who were recruited for the purpose of the study for one short session only as part of counsellor training.

Whittaker also assessed whether different types of supervision – ie individual supervision, group supervision and live supervision – have varying effects on supervisees' anxiety and self-efficacy. Her findings concluded that clinical supervision increased supervisees' self-efficacy. Different participants in the interviews described varying supervision methods that increased supervisees' anxiety and self-efficacy, but the results were unable to establish a pattern in terms of what type of supervision or techniques increase or decrease anxiety and self-efficacy to the greatest extent. No studies examined peer group supervision on self-efficacy; therefore, no conclusion was reached as to the different types of supervision's effect on self-efficacy other than individual supervision. However, given those limitations, clinical supervision was found to have a medium Effect Size on supervisees' anxiety and a large Effect Size on their self-efficacy. She concluded that "supervisors need to be trained to recognise low self-efficacy of supervisees and learn effective methods to address these concerns" (Whittaker, 2004: 51).

Section 3: Methodology

3.1 Structure of review tasks

This systematic review is a modified version of a Cochrane Review (Higgins & Green, 2005) conducted with the use of the EPPI-Reviewer software (see pg 9 for more details). Preparation for conducting the review involves careful consideration of the scope and focus of the work. The process is rigorous and well defined, while maintaining a practical perspective. Tasks include:

- identifying the precise definition of a systematic review question
- determining the criteria for study selection
- creating and using a protocol for recording the search process
- creating and using a protocol for reviewing, recording and rating the relevance and methodological quality of each study
- analysing the data collected from the studies as appropriate
- ensuring that the analysis takes account of the different training systems and models of counselling and psychotherapy used internationally that may influence the outcome
- writing a report and wider dissemination to a range of stakeholders, including policy makers, practitioners and researchers.

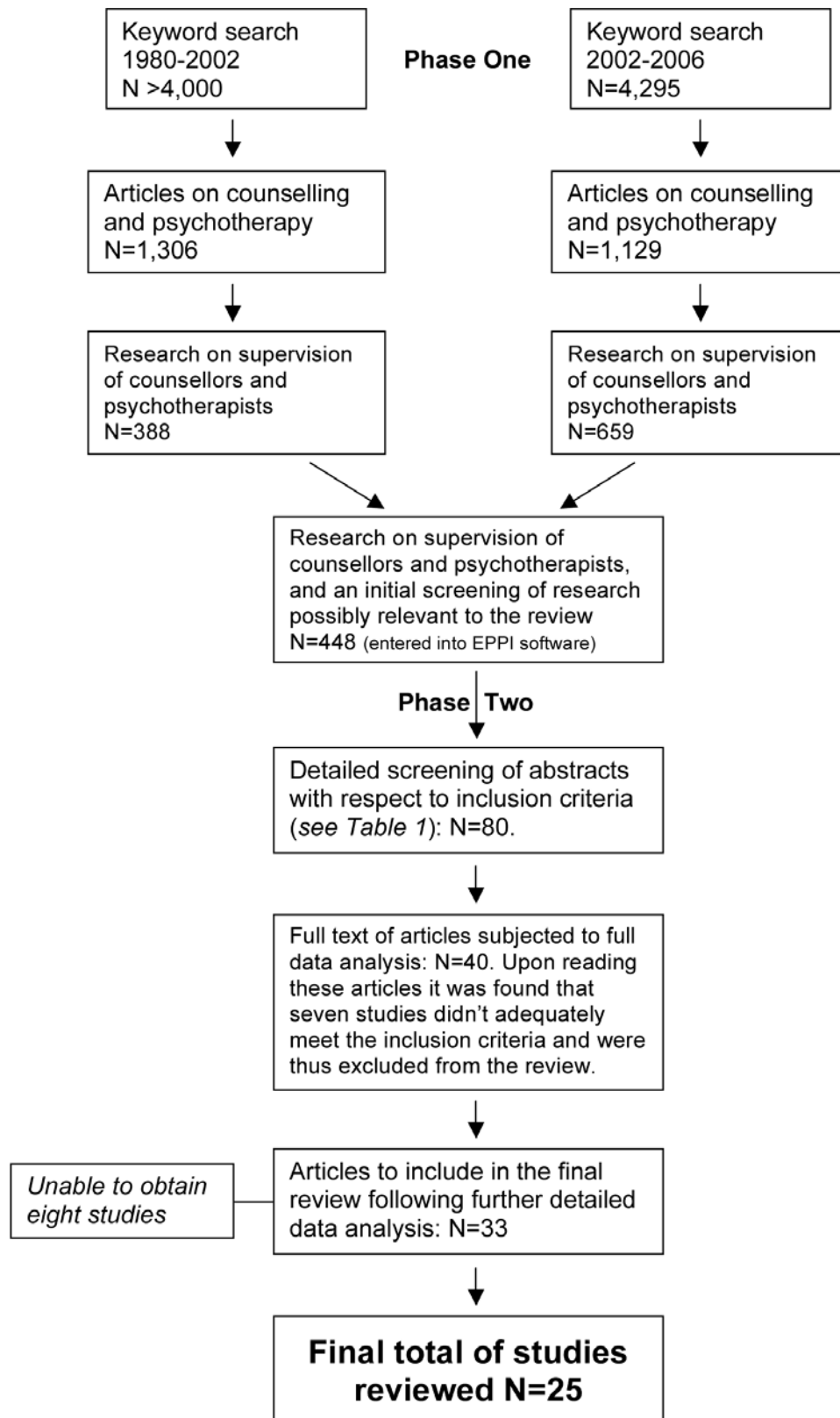
This review was carried out in two main phases. Phase One involved the scoping of the literature on supervision in order to identify the parameters of the research question and the main exclusion/inclusion criteria. The previous scoping search undertaken by Wheeler (2003) helped inform this process. The literature search of relevant studies/articles was also completed in this phase and, initially, abstracts of articles were screened to determine whether they met the inclusion criteria. The literature search previously carried out by Wheeler (2003) was used for articles from 1980–2002 and an updated search was carried out for articles from 2002 to early 2006.

Figure 1 (see over page) provides an overview of the literature search and retrieval process during this Phase.

Phase Two commenced with the acquisition of the full text of articles that looked from their abstracts as though they would meet the inclusion criteria. Detailed scrutiny of these articles led to the exclusion of 50% (n=40) of those identified in Phase One, because they didn't meet the criteria. The remaining articles available were subjected to a detailed data extraction procedure, during which process seven further articles were excluded – this left 33 studies to be included in the review. From these 33 studies, eight were unobtainable and so the final total of studies in this systematic review is 25 (see pg 11), section 4.1.1, 'An overview of included studies', for a full list of all these reviewed studies).

Figure 1: Flowchart

Phase One and Phase Two: an overview of literature search/retrieval, initial screening of studies and data extraction



3.2 Phase One: scoping and retrieval of relevant studies

3.2.1 Scope of the review: defining inclusion and exclusion criteria

This review examined the evidence related to the impact of supervision on counsellors and psychotherapists, their practice and their clients. It was decided to include published and unpublished work in English from 1980. The review sought to find and analyse empirical research (both quantitative and qualitative) that produced data that had been subject to systematic analysis; discursive articles and case studies were excluded.

Initially, a sample of 10 available articles that seemed to examine some impact of supervision on the supervisee were read in detail to help inform the identification of appropriate inclusion and exclusion criteria.

Prior to commencing the research, it was hypothesised that studies might be found that would focus on four aspects of supervisee functioning:

- clinical competence
- ethical practice
- personal development and well-being
- organisational competence: the capacity to understand and work with organisational dynamics.

It was intended to categorise articles according to these descriptors. It was considered important to include studies that had an objective measurement of the impact of supervision on the supervisee. However, knowledge of the literature suggested that this would severely limit the number of identified studies and the criteria were extended to include self-report. However, self-report of satisfaction with supervision was specifically excluded, as satisfaction would not necessarily correlate with any change, ie development of competence or other benefits of supervision. Studies were included if a valid and reliable instrument was used to quantify the impact of supervision on the supervisee or if there was a methodologically rigorous qualitative analysis of the impact. This requires that the method of analysis is clearly articulated in the article and complies with this description.

For this review it was decided that the supervisees must be counsellors or psychotherapists (or other professionals who have had a substantial training as counsellors or psychotherapists) and who were specifically engaged in a counselling role with clients. Hence psychiatrists, psychiatric nurses, occupational therapists and other health professionals having supervision were excluded. There are a considerable number of studies that report on supervision of family therapy. However, family therapy often involves immediate supervision from a supervisor behind a screen, with an audio link or even in the therapy room, which was considered to be substantially different from one-to-one, face-to-face supervision that relies on recall or tape recordings of sessions. Hence such studies were excluded. The supervision of art and music therapy was also excluded, mainly because training in these approaches is somewhat different from training in mainstream counselling and psychotherapy. For this reason, models of supervision will often differ.

Supervision often happens in pairs or groups, particularly during training, but group supervision was only included if it related to therapeutic work with individuals or couples. Group and individual supervision focusing on groupwork studies were therefore excluded. Studies were also only included if

they specifically focused on a supervision intervention that related to work with real clients. So supervision of role-play sessions for the purpose of skill development was excluded. A lot of supervision research is undertaken in the USA and, thus, most studies relate to the supervision of trainees. Careful attention was paid, therefore, to recording whether the supervision related to trainees or experienced practitioners, but both were included in the review.

Technology is finding a role in counselling and psychotherapy, and supervision is no exception. When online or telephone supervision of counselling and psychotherapy met the other inclusion criteria for the review, they were incorporated. One-to-one peer supervision and peer group supervision were excluded, as they would be unlikely to have reliable estimates of supervision quality. If mainstream definitions for supervision are considered, the supervisor is usually more experienced than the supervisee, taking responsibility for all aspects of the supervisory process, eg role distinction and addressing ethical issues. However, in peer supervision the roles between supervisee and supervisor are not always sufficiently distinguished. Furthermore, given that most supervision research has been undertaken in the US, most peer supervisors will be fellow trainees. Common expectations of supervision would not usually include trainees acting as supervisors. Level of professional experience, accountability and clear role definition are not adequately taken into account. Consequently, this kind of supervision is less likely to be comparable with mainstream clinical supervision.

Table 1 (see over page) provides an overview of the exclusion criteria and the number of studies excluded based on this. Some studies may have been excluded on more than one criteria.

3.2.2 Search Strategy

The 2003 scoping search on supervision (Wheeler, 2003) was used as the main source of articles prior to 2002. The same search strategy was used for obtaining more recent studies from 2002 to early 2006. Key words relating to counselling and psychotherapy were truncated for literature searching. These included *supervis**, *counsel**, *therap** and *psychotherap** and were used in searching electronic bibliographic databases (including ERIC, PsychInfo, Medline, EMBASE, WorldCat Dissertations, IBBS and the Cochrane Database) for studies from 2002 to early 2006. Information was recorded about all studies from 1980 to 2006 that might meet the inclusion criteria after a superficial inspection of abstracts (see Figure One for an overview of the literature search and retrieval of studies).

Studies were initially included or excluded after reading the titles and abstracts of possible articles. Relevant journals were hand-searched for articles and bibliographies of books on supervision published since 2002 were consulted. All relevant citations were downloaded into the Endnote bibliographic referencing system; the initial screening of citations was conducted using Endnote. When this initial screening was complete, 448 references were uploaded into the EPPI-Reviewer software for a more detailed screening process.

3.2.3 EPPI-Reviewer

In 1995, the Evidence for Policy and Practice Information and Coordinating (EPPI) Centre was firmly established. This occurred when the Department of Health commissioned a series of reviews in the area of health promotion to mirror the work of Cochrane in the field of non-clinical health issues. In 2000, the EPPI-Centre broadened its scope and started to undertake reviews in the field of education.

Table 1: Exclusion criteria for the screening of articles/studies

Criteria name	Criteria details	No of final studies excluded based
EXCLUDE supervision before 1980	Search limited until after 1980.	N/A
EXCLUDE self-report satisfaction	Exclude studies when self-report of satisfaction with supervision or the supervisor is the ONLY data reported.	50
EXCLUDE role-play or simulated therapy or supervision	The counselling must be with real clients who have consented to treatment; role play or simulated therapy or supervision is excluded.	7
EXCLUDE family therapists, group therapy, hypnotherapy	Exclude family therapists, group therapy, hypnotherapies, but include couple counsellors.	66
EXCLUDE psychiatrists, nurses, GPs, social workers	Psychiatrists, nurses, GPs, social workers, teachers etc can only be included if they are clearly defined as being in a formal contracted counselling role.	10
EXCLUDE studies that do not report on any aspect of impact	Only include studies that examine some aspect of the IMPACT of supervision on the supervisee, their practice and their clients.	233
EXCLUDE supervision of groupwork and peer supervision	All forms of peer supervision are excluded. Supervision undertaken in pairs or groups is included. Supervision of group work is excluded.	6
EXCLUDE art therapy and music therapy	Other forms of therapy, eg art therapy and music therapy, are to be excluded. However, online counselling can be included.	0
EXCLUDE Discursive articles	Exclude articles that do not include any systematic collection of data related to supervisees.	46
EXCLUDE unavailable	Studies that were listed for data-extraction but were unavailable.	8

The methodology, expertise and tools developed within the EPPI-Centre have resulted in reviews being undertaken in many fields, appraising and synthesising a broad range evidence for all research questions and thus potentially to include all types of study (EPPI, 2006b). Part of the developments of the EPPI-Centre has been the advance of software to aid the process of undertaking a systematic review – this is known as EPPI-Reviewer. EPPI-Reviewer provides a comprehensive online software system for storing all information and citations for a systematic review, and for analysing and synthesising data (EPPI-Reviewer software). This was the first time that BACP has used EPPI-Reviewer for any of its scoping reviews, and thus this review acted as a pilot on the use of EPPI for BACP reviews. Other reviews being undertaken by BACP are now using the EPPI-Reviewer. This means that all references/data extraction from BACP's scoping reviews will be in one central searchable database, and other reviewers will be able to have immediate access to previous and ongoing reviews (including access to all stages of screening and data extraction information).

3.2.4 Reviewing, recording and rating relevance of methodology of each study

EPPI-Reviewer was programmed with the inclusion and exclusion criteria, and the articles were reviewed against these criteria. A data-extraction pro-forma was designed and programmed into EPPI-Reviewer for recording information from included studies (see Appendix for a copy of the form). This data-extraction procedure addressed pertinent

questions, such as type of study, along with reliability and validity issues such as: research methodology, subjects, sample size and research bias. As part of this data-extraction process, 11 potential supervision impact categories were identified. These impact categories were assessment skill, self-efficacy, self-confidence, competence, case conceptualisation, counselling skills, ethical practice, decision making, knowledge, outcome for patient and others (as they emerged from literature). During the data-extraction phase, studies that were initially included that were found not to meet the inclusion criteria and were excluded from the review. The categories that were expected in the first instance did not relate closely to the studies that were found and hence were ignored.

A team of four experienced counselling researchers participated in the data-screening and extraction process. Initially, one reviewer assessed abstracts. Any abstracts that were ambiguous were either discussed by two reviewers or left in until the full text article had been obtained. This provided more detailed information, enabling a robust decision to be made about inclusion or exclusion. Approximately half of the papers in the review were subjected to detailed data extraction by two independent reviewers and these were cross-checked. The remaining articles were data extracted by one person and were not cross checked. Limited time and resources did not allow for all articles to be double extracted and cross checked. However, all queries were referred to the two principal reviewers, who also corrected the data extraction when gaps were noted.

Section 4: Results

The research literature examined in this section of the report focuses on the results from the final list of 25 studies that met the inclusion criteria and, thus, were deemed relevant to the research question 'what is the impact of clinical supervision on counsellors, therapists, their practice and their clients?' This section starts with identification of the 25 studies, along with an overview of the supervision impact categories that have been identified in relation to these studies. The methodological approaches taken in the studies (including details of the instruments used), together with detailed information about the supervisees, supervisors and clients who took part in the studies, are reported. The relevant results from these studies are discussed, along with some commentary on the quality of the evidence from each study.

4.1 An overview of the titles of studies included and excluded

4.1.1 An overview of titles of included studies

The research literature examined in this section is taken from the 25 articles/studies that were included in the final review, following the systematic retrieval and inclusion process as described earlier. These studies are listed in Table 2 and full abstracts of each of these articles/studies are listed in Section 7.1: 'References'.

4.1.2 An overview of titles of studies unobtainable and excluded

'Section 7: Research Articles' includes a list of the titles and abstracts of studies that were initially included in the review but were unobtainable and thus data extraction was not completed for them. Had they been available, they may or may not have been included in the review.

Table 2: A list of the 25 articles/studies included in this review

Bambling, M., King, R., Raue, P., Schweitzewr, & Lambert, W. (2006). Clinical supervision: Its influence on client-rated working alliance and client symptom reduction in the brief treatment of major depression. <i>Psychotherapy Research</i> , 16 (3): 317–331.	Milne, D. L., Pilkington, J., Gracie, J., & James, I. (2003). Transferring skills from supervision to therapy: A qualitative and quantitative <i>N</i> =1 analysis. <i>Behavioural and Cognitive Psychotherapy</i> , 31: 193–202.
Borders, L. D. (1990). Developmental changes during supervisees' first practicum. <i>Clinical Supervisor</i> , 8 (2): 157–167.	Ogren, M. L., & Jonsson, C. O. (2003). Psychotherapeutic skill following group supervision according to supervisees and supervisors. <i>The Clinical Supervisor</i> , 22 (1): 35–58.
Carey, J. C. Williams, K. S. & Wells, M. (1988). Relationships between Dimensions of Supervisors' Influence and Counselor Trainees' Performance. <i>Counselor Education and Supervision</i> , 28: 130–139.	Ogren, M. L., Jonsson, C. O., & Sundin, E. C. (2005). Group supervision in psychotherapy. The relationship between focus, group climate and perceived attained skill. <i>Journal of Clinical Psychology</i> , 61 (4): 373–388.
Cashwell, T. H. & K. Dooley (2001). The impact of supervision on counselor self-efficacy. <i>Clinical Supervisor Special Issue</i> , 20 (1): 39–47.	Ossana, S. M. (1990). <i>The impact of supervision on trainee self-efficacy and counseling ability: Testing a developmental framework</i> . PhD dissertation: University of Maryland College, USA.
Couchon, W. D. & Bernard, J. M. (1984). Effects of timing of supervision on supervisor and counselor performance. <i>The Clinical Supervisor</i> , 2 (3): 3–20.	Patton, M. J., & Kivlighan, D. M. (1997). Relevance of the supervisory alliance to the counseling alliance and to treatment adherence in counselor training. <i>Journal of Counseling Psychology</i> , 44 (1): 108–115.
Dodenhoff, J. T. (1981). Interpersonal attraction and direct-indirect supervisor influence as predictors of counsellor trainee effectiveness. <i>Journal of Counseling Psychology</i> , 28 (1): 47–52.	Raichelson, S. H., Herron, W. G., Primavera, L. H., & Ramirez, S. M. (1997). Incidence and effects of parallel process in psychotherapy supervision. <i>The Clinical Supervisor</i> , 15 (2): 37–48.
Earley, T. M. (2004). <i>Counselor supervision post-licensure: Relationship of clinical supervision type and frequency to supervision anxiety and counselor self-efficacy</i> . PhD dissertation: Capella University, USA.	Steinhelber, J., Patterson, V., Cliffe, K., & LeGoullon, M. (1984). An investigation of some relationships between psychotherapy supervision and patient change. <i>Journal of Clinical Psychology</i> , 40 (3): 1346–1353.
Efstation, J. F., Patton, M. J., & Kardash, M. (1990). Measuring the working alliance in counsellor supervision. <i>Journal of Counseling Psychology</i> , 37 (3): 322–329.	Strozier, A. L., Kivlighan, D. M., & Thoreson, R. W. (1993). Supervisor intentions, supervisee reactions and helpfulness: A case study of the process of supervision. <i>Professional Psychology: Research and Practice</i> , 24 (1): 13–19.
Guest, P. D., & L. E. Beutler (1988). Impact of Psychotherapy Supervision on Therapist Orientation and Values. <i>Journal of Consulting & Clinical Psychology</i> , 56 (5): 653–58.	Tryon, G. S. (1996). Supervisee Development during the Practicum Year. <i>Counselor Education & Supervision</i> , 35 (4): 287–94.
Jiménez Andújar, D. J. (2002). <i>Experiences of personal growth of Masters-level supervisees within the supervisory process</i> . PhD dissertation: University of New Orleans.	Vallance, K. (2004). Exploring counsellor perceptions of the impact of counselling supervision on clients. <i>British Journal of Guidance and Counselling</i> , 32 (4): 559–574.
Jumper, S. A. (1999). <i>Immediate feedback using the 'bug-in-the-ear' in counselor training: Implications for counseling self-efficacy, trainee anxiety and skill development</i> . PhD Dissertation: University of North Dakota.	Worthen V., & McNeill, B. W. (1996). A phenomenological investigation of 'good' supervision events. <i>Journal of Counseling Psychology</i> , 43 (1): 25–34.
Ladany, N., Ellis, M. V., & Friedlander, M. L. (1999). The supervisory working alliance, trainee self-efficacy and satisfaction. <i>Journal of Counseling and Development</i> , 77: 447–455.	Worthington, E. L. (1984). Empirical investigation of supervision of counselors as they gain experience. <i>Journal of Counseling Psychology</i> , 31 (1): 63–75.
Leherman-Waterman, D., & Ladany, N. (2001). Development and validation of the evaluation process within supervision inventory. <i>Journal of Counseling Psychology</i> , 48 (2): 168–177.	

'Section 8: References' lists the studies/articles that were excluded during Phase Two. Listing these references offers transparency of studies excluded during the data-extraction process. Details of all other excluded references are available directly from the authors. However, many of the 8,000 and more abstracts screened did not bear relevance to either research or the supervision of counselling and psychotherapy. They would serve no other use in being examined other than for transparency purposes. There were 448 references of more relevance, as they reported supervision research related to counselling and psychotherapy although beyond the remit of this review.

4.2 Supervision impact categories

As part of the data-extraction process, supervision impact categories were identified. Table 3 gives a general overview of the frequency of these across the studies examined (note that some studies are relevant to more than one impact category and thus the frequencies shown are representative of this). The table also offers an early indication of the areas in which the most supervision research has been undertaken. The majority of studies report on the development of the supervisee's skill and self-efficacy.

The 'Other' category included the following:

Upon completion of data extraction, the studies were organised into five overarching impact categories. These are:

1. **Self-awareness** (three studies in total)
2. **Skill development** (nine studies in total)
3. **Self-efficacy** (seven studies in total)
4. **Transfer of understanding from supervision to therapy** (five studies in total)
5. **Client outcome** (one study in total)

Table 3: The frequency of supervision in initial impact categories

Supervision impact examined	Number of studies
self-efficacy	8
self-confidence	3
competence	5
case conceptualisation	1
general skills	11
assessment skill	1
ethical practice	1
decision making	1
knowledge	2
outcome for patient	3
Others including: Self-other awareness Trainee/supervisee anxiety Values and theoretical orientation Personal growth Support/encouragement/challenge Timing and frequency of supervision Working alliance Professional development	18

Table 4 provides an overview of those studies in each of these final supervision impact categories. Some of the studies may be relevant to more than one category. However, for ease of interpretation, the studies are organised into the category that they best represent.

4.3 Methodological approaches and measures used in the studies

The studies in the review used a range of methodological approaches, including qualitative, quantitative and mixed methods. A total of 16 studies used only quantitative measures, four used only qualitative methods and six used mixed methods (Table 5 lists which of the studies used which general methodological approach). Table 6 gives a more detailed overview of the range of research methods used in the studies – methods used included self-completion reports, video recordings, open-ended questionnaires, semi-structured interviews and observations of the supervisee. The most frequent method used to collect data was questionnaires with supervisees. Table 7 identifies the standardised questionnaires used in these studies.

In addition to the instruments listed in Table 7, the range of other methods used in each study can be identified in Table 11 in 'Section 4.5: Research Evidence'.

4.4 The supervisees, supervisors and clients in the studies

The information provided about supervisees, their clients and supervisors taking part in the studies is varied. Some studies included measures that were completed by supervisees and supervisors, and these studies tended to provide information about supervisors as well as supervisees. A few collected data and reported on clients. The studies that only gathered data from supervisees seemed to omit important information about the supervisors working with the supervisees. Detailed information about the context and possible variables in supervision are omitted from many studies (eg ethnicity, theoretical orientation of the supervisees and supervisor, and years of experience of supervisors).

4.4.1 The supervisees

Information provided about the supervisees offers more detail than that for the supervisors. In those studies that reported the information, age and gender are reported in 18 studies. Ages between 27 and 35 years are the most frequent, with the full age range across studies being 22 to 73 years old. Fifteen studies had both male and female subjects but more females than male supervisees participated in the studies. Two studies had only females in them and 5 studies did not report on the gender of the supervisees. Ethnicity of the supervisees seemed to be overlooked in more than half of the studies, but seven had subjects that were of mixed ethnicity. In these studies, at least 85% of participants were white Caucasian American. Asian-American, Latina, African-American, black and Hispanic participants were referred to in a few studies, while one study included all white British subjects (Vallance, 2004).

Supervisees were trainees in over two-thirds of the studies (18 studies). This figure is not surprising given that most of the studies are undertaken in the USA. The theoretical orientation of the supervisee and of the therapeutic practice did not yield any clear trends – between 40 and 50% of studies did not provide any information about theoretical orientation. When information was provided, it seemed that all the main theoretical orientations were represented. However, when

Table 4: An overview of supervision impacts for each study

Self-awareness	Skill development	Self-efficacy	Transfer of understanding from supervision to therapy	Client outcome
Jiménez Andújar (2002)	Borders (1990)	Cashwell, & Dooley (2001)	Couchon & Bernard (1984)	Bambling et al., (2006)
Raichelson, Herron, Primavera, & Ramirez (1997)	Carey, (1988)	Earley (2004)	Guest & Beutler (1988)	
Tyron (1996)	Dodenhoof (1981)	Efstation, Patton, & Kardash (1990)	Milne, Pilkington, Gracie & James (2003)	
	Ogren & Jonsson (2003)	Jumper (1999)	Patton & Kivlighan (1997)	
	Ogren, Jonsson & Sundin (2005)	Ladany, Ellis, & Friedlander (1999)	Steinhelber, Patterson, Cliffe, & Legoullon (1984)	
	Strozier, Kivlighan, & Thoreson (1993)	Lehrman-Waterman & Ladany (2001)		
	Vallance (2004)	Ossana (1991)		
	Worthen & McNeill (1996)			
	Worthington (1984)			

stated, psychodynamic seemed to be more prominent than other orientations. Finally, information about whether the supervisees had been given any training on using supervision was only stated in one study (Jumper, 1999).

4.4.2 The supervisors

Demographic information for supervisors is reported less than for supervisees. Data related to supervisors were not included in many of the studies, an important omission that affected the quality of many of the studies. The age of supervisors is not stated in 18 studies, but, when stated, the age range is 22 to 60 years old. The gender of supervisors is not stated in 13 of the studies, but, when stated, it tended to be a mix of male and female. Ethnicity of the supervisors was not stated in 20 of the studies. Only two studies (Lehrman-Waterman & Ladany, 2001; Ossana, 1991) reported mixed ethnicity, where 85% of participants were white Caucasian American. Other represented ethnicity included Asian American, Black and Hispanic. One study reported that all subjects were European Americans (Patton & Kivlighan, 1997).

The supervisors' years of experience as therapists and supervisors were not stated in 16 of the studies. In terms of theoretical orientation of the supervisor, the theoretical model of therapy was not stated in approximately 45% of studies. When stated, psychodynamic seemed to be more prominent – this was identified in four studies. The theoretical model of supervision was not stated in 17 of the studies. When stated, a variety of theoretical models were identified (including interpersonal, psychodynamic and CBT), but this only equated with, at most, two of the studies. As these studies are evaluating the impact of supervision, the omission of these potentially highly relevant facts and variables seems very surprising.

4.4.3 The clients

As is evident from the overview of clients in the studies, most studies do not report any information about the clients with whom the counsellors or psychotherapists were working. Most studies concentrated on the experiences of the

supervisees. The study by Bambling et al (2006) describes in detail the presenting problems of clients, because it examines the impact of supervision on client symptom reduction in the brief treatment of major depression. Steinhelber (1984) also gives a breakdown of client diagnosis. Some of the other studies offer limited client information, for example, Jumper (1999), Patton & Kivlighan (1997) and Strozier et al (1993). Information about clients seen by supervisees in supervision research is of paramount importance and reflects one of the methodological issues of supervision research that needs further exploration and transparency in studies.

4.5 Quality of the evidence

4.5.1 Quality of the evidence and study limitations

'The hierarchy of evidence' is often used in systematic reviews as a way of judging the scientific rigour of studies, as it reflects the relative weight given to different types of research evidence. When using this type of weighting, the evidence at the top of the hierarchy (systematic reviews and meta-analyses) is seen as the 'best quality evidence' – ie the most scientifically rigorous. In this review, studies were not ranked against this hierarchy, as philosophically it didn't fully match the pluralistic research stance of the authors. Furthermore, there are disadvantages for the psychological therapies in maintaining this rigid hierarchy. For example, Rowland (2007) points out that because there is a lack of systematic review and Randomised Control Evidence (RCT) evidence for the psychological therapies, a hierarchy of evidence maintains a very narrow evidence base.

It is recognised, however, that knowing the type of study will be important for readers in order for them to make their own judgments about how the research was conducted. Hence, the methodological approach of each study is stated (see Table 8). However, knowing the methodological type of the study doesn't guarantee quality, as any study could still have a range of methodological flaws. For example, a well-designed cohort study could be better quality than a poorly designed RCT.

Table 5: An overview of general methodological approaches to studies

Quantitative studies	Mixed-methods studies	Qualitative studies
Bambling, King, Raue, Schweitzer & Lambert (2006) <i>Randomised Controlled Trial (RCT)</i>	Couchon & Bernard (1984) <i>Questionnaires and videotape recordings</i>	Jiménez Andújar (2002) <i>Cross-sectional semi-structured interviews</i>
Borders, L. D. (1990) <i>Longitudinal questionnaire</i>	Guest & Beutler (1988) <i>Questionnaires and cross-sectional survey</i>	Milne, Pilkington, Gracie & James (2003) <i>Single case study design</i>
Carey, Williams & Wells (1988) <i>Cross-sectional questionnaire</i>	Patton & Kivlighan (1997) <i>Questionnaires, observations and videotapes</i>	Vallance (2004) <i>Qualitative questionnaires and semi-structure interviews</i>
Cashwell & Dooley (2001) <i>Cross-sectional questionnaire</i>	Raichelson, Herron, Primavera, & Ramirez (1997) <i>Cross-sectional survey</i>	Worthen & McNeill (1996) <i>Semi-structured interviews</i>
Dodenhoff (1981) <i>Longitudinal questionnaire (pre, mid and post)</i>	Strozier, Kivlighan & Thoreson (1993) <i>Questionnaires and videotape ratings</i>	
Earley (2004) <i>Questionnaires and cross-sectional survey</i>		
Efstation, Patton & Kardash (1990) <i>Quantitative construction of a measure</i>		
Jumper (1999) <i>Longitudinal quantitative (pre, mid and post)</i>		
Ladany, Ellis, & Friedlander (1999) <i>Longitudinal quantitative</i>		
Lehrman-Waterman & Ladany (2001) <i>Cross-sectional survey: construction of a measure</i>		
Ogren, Jonsson & Sundin (2005) <i>Quantitative cross-sectional questionnaire</i>		
Ogren & Jonsson (2003) <i>Longitudinal questionnaire (pre and post)</i>		
Ossana (1991) <i>Longitudinal questionnaires</i>		
Steinheiber, Patterson, Cliffe & Legouillon (1984) <i>Longitudinal questionnaires</i>		
Tyron (1996) <i>Longitudinal questionnaire (pre, mid and post)</i>		
Worthington (1984) <i>Cross-sectional survey</i>		

It is important to quality assess studies further in order to ascertain how well they have been designed, whether the study design was rigorous, whether the results are trustworthy and whether conclusions can be viewed with confidence. Many factors determine whether a study can be deemed reliable/trustworthy. Potential sources of bias need to be assessed to identify whether the results of the study need to be treated with caution – often the findings of a study may support or disprove the hypothesis, but methodological weaknesses mean the quality of the evidence is questionable and thus caution is needed when interpreting results.

In this review, a quality threshold for the inclusion of studies was not predetermined. However, all studies had to report on empirical research and those that hadn't were naturally excluded. The criteria for assessing the standard of qualitative and quantitative research differ. The Downs and Black Rating Scale (Deeks et al., 2003) and the Critical Appraisal Skills Programme (CASP) 10 questions (Milton Keynes Primary Care Trust, 2002) are examples of tools that can be used to help rate studies or make sense of qualitative research to support the process of assessing the quality of evidence. These tools ask key methodological questions that enable an assessment

Table 6: An overview of types of methods used to collect data

Method used to collect data	Number of studies
Interviews – supervisee	4
Interviews – supervisor	1
Observation – of supervisee	1
Self-completion report/diary – supervisee	2
Self-completion report/diary – supervisor	1
Standardised q'aire – supervisee (specify q'aire)	15
Standardised q'aire – supervisor (specify q'aire)	5
Other q'aire – supervisee (specify q'aire)	12
Other q'aire – supervisor (specify q'aire)	5
Video recording – supervisor	3
Video recording – supervisee	3
Video recording – client	1
Standardised q'aire – client (specify q'aire)	3
Other q'aire – client (specify)	1
Others (specify)	5

of the studies' limitations. In undertaking this review, we had restricted resources, but we still wanted to indicate limitations of studies and offer some kind of initial quality rating. We provided, therefore, a holistic judgement of the study quality, in which we categorised each study as either 'average', 'good' or 'excellent'.

Average studies were those asking a relevant research question related to the impact of supervision and had some useful related results. However, confidence in the results was limited for a range of reasons. This could have been that the research question although relevant might not have been clear, the methodology could have been convoluted, weaknesses could have been clearly evident in the methodology and/or it was a relatively small sample size. Good studies, although with some limitations, seemed to offer some confidence in the results and had tried to overcome some of the inherent limitations in supervision research. And excellent studies were those that represented clear and well thought-out methodology, whereby the research design had attempted to address the variety of inherent problems addressed with supervision impact research. Overall, in excellent studies the results from the study could be confidently linked with supervision outcome.

The holistic judgement made on the quality of studies was drawn from the data extraction. Many of the data extraction questions provided information related to quality issues of research methodology. For example, items such as clarity of question, sample procedures, socio-demographics of participants, allocation of control comparison groups, study biases, factors favourable to developing and delivering the intervention, study limitations, inconsistencies in reporting, and the relevance of findings for the future development of supervision are relevant to quality judgements. It was clear from reading studies whether the research design was appropriate to the aims of the research question, what biases

in data collection were evident, whether the research process had been rigorous and whether the findings were fairly reported.

Detail of the research design provides insight into what the biases and weaknesses of studies might be. A range of limitations is identified across studies. These are discussed in detail for each study in Section 4, which reports on the research evidence. However, some of the methodological problems of studies are discussed below. For some studies it proved difficult to fully assess the quality of the study, as the design and approach were complex and often convoluted. Thus, the rating scores given for the studies in this review should only be used as an initial indication of confidence, as they are not based on a thorough systematic analysis of quality. Furthermore, some of the studies categorised in the 'average' category will be better designed than other studies also in this category. However, collectively these quality ratings give a general indication as to the overall quality of supervision outcome research.

4.5.2 Examples of methodological problems

Some authors identified some of the limitations of their research. For example, in an RCT study, Bambling et al. (2006) noted that it was not possible to separate the supervision effect from any pre-treatment session and therapist allegiance effects. They also noted that, although providing a moderate sample size, total power was insufficient to ensure that Type II errors could not occur. Also the principal researcher in this study undertook a significant amount of the supervision across both of the study conditions, and researcher bias may have affected results.

In a longitudinal study, Borders (1990) identified limitations associated with the relatively small number of participants from one counselling program (and the reliance on one self-report measure of developmental change). Earley (2004) also noted similar limitations in a cross-sectional survey study, including the fact that the study was a convenience sample and that it was an ex-post factor, correlation study that could not establish a relationship among the variables. Guest and Beutler (1988) pointed out that, in their cross-sectional survey study, the sample size was also limited and large numbers of variables were examined. Lehman-Waterman and Ladany (2001) noted that limitations in their cross-sectional survey study included a reliance on a single method approach and the problem that participants self-selected to complete the study survey. Other limitations included lack of a control group, with no comparison to a no-supervision condition. Other methodological issues were noted – Strozier et al. (1993) pointed out that, because they had used a case study design, their results are best seen as suggestive and heuristic, and the implications that can be drawn from the results of the study should be further investigated with group designs.

Along with these limitations identified by the authors, the review process highlighted further limitations for studies. Some studies failed to provide pertinent information about the context of the supervision. Often no clear definition of supervision was given, information about the supervisee, supervisor and clients was limited, and measurements were only taken at one point in time. It was common for the socio-demographic variables in the study not to be taken into account. For example, in a study by Patton & Kivilghan (1997), the raters were all European American, whereas the participants were a mixture of African American and European American, but this was not highlighted. Tracey and Sherry (1993) provided a comparison between counsellor/client and supervisor/supervisee interactions, and yet there were no details of the clients in the study.

Table 7: Instruments used in the studies reviewed

Questionnaire	Study
Self Evaluation Scale (Buckley et al, 1981, 1982). Also modified Swedish version	Ogren, Jonsson & Sundin (2005) Ogren & Jonsson (2003)
Self-efficacy Inventory (Friedlander & Snyder, 1983)	Ladany, Ellis & Friedlander (1999)
Supervisee Levels Questionnaire (SLQ) (McNeill, Stoltenberg & Pierce, 1985)	Borders (1990), Tyron (1990)
Global Assessment of Functioning (Endicott et al, 1976)	Steinhelber et al (1984)
Counseling Evaluation Inventory (Linden, Stone and Shertzer, 1965)	Couchon & Bernard (1984)
Counseling Self-estimate Inventory (Larson et al, 1992)	Cashwell & Dooley (2001) Jumper (1999)
Session Evaluation Questionnaire (Stiles & Snow, 1984)	Strozier, Kivlighan & Thoreson (1993)
Supervisor Intentions List (modification of Hill & O'Grady's Therapist Intention List, 1985)	Strozier, Kivlighan & Thoreson (1993)
Supervisees Reactions System (researchers' modification of the Client Reactions System (Hill et al, 1988)	Strozier, Kivlighan & Thoreson (1993)
Counselor Rating Form (Barak & LaCrosse, 1975)	Dodenhoff (1981)
Rating Scale for Outcome (Storrow, 1960)	Dodenhoff (1981)
Counselor Evaluation Rating Scale – CERS (Myrick & Kelly, 1971)	Dodenhoff (1981) Carey et al (1988) Jumper (1999)
Theoretical Orientation Questionnaire (Sundland, 1977a, 1977b)	Guest & Beutler (1988)
Trainee Personal Reaction Scale Revised (Holloway & Wampold, 1984)	Ladany, Ellis & Friedlander (1999)
Value Survey (Rokeach, 1973, 1979)	Guest & Beutler (1988)
Supervisor Working Alliance Inventory (Efstation et al, 1990)	Patton & Kivlighan (1997) Efstation et al (1990)
Supervisory Styles Inventory (Friendlander & Ward, 1984)	Efstation et al (1990)
Working Alliance Inventory Trainee Version (Bahrck, 1990)	Ladany, Ellis & Friedlander (1999)
Working Alliance Inventory (Horvarth & Greenberg, 1989)	Patton & Kilighan (1997) Bambling et al (2006)
Locus of Control (Rotter, 1966)	Guest & Beutler (1988)
Helpfulness Rating Scale (Elliott, 1985)	Strozier, Kivlighan & Thoreson (1993)
TAC – Topics and Climate Inventory	Ogren et al (2005)
Supervisor Rating Scale (SRF) (Heppener & Handley, 1981, 1982)	Carey et al (1988)
State-Trait Anxiety Scale (STAI) (Spielberger, 1970)	Jumper (1999) Earley (2004)
Challenging Skills Rating Form (CSRf) (Johnson et al 1989)	Jumper (1999)
Trainee Value of Cues Scale (TVCS) (McClure & Vriend (1976)	Jumper (1999)
The Counsellor Self-efficacy Scale (CSES) (Melcart et al, 1996)	Earley (2004)
The Supervision Questionnaire (SQ) and the Revised Supervision Questionnaire (SQ-R) (Worthington & Roehlke, 1979)	Worthington (1987)
The Social Skills Inventory	Bambling et al (2006)
Supervision Focus Adherence Scale (SFAS)	Bambling et al (2006)
Beck Depression Inventory	Bambling et al (2006)
Supervision Level Scale (SLS) (Wiley & Ray, 1986)	Ossana (1990)
Counselor Ability Scale: CAS-A: Supervisee counseling ability and CAS-E: Supervisee self-efficacy	Ossana (1990)

The following is an overview of these and other general methodological weaknesses that are evident across many of the studies:

- convoluted questions and discussions
- the majority of studies were on trainees; this makes it difficult to differentiate supervision impact from that of other training
- the reliance on self-report measures
- the reliance on single measures
- limited information on supervisees, supervisors and clients (most of the time there was no information on clients)
- no definition of supervision
- lack of clarity about the role of the researcher
- small sample sizes
- measures taken at only one point
- methods of data collection and analysis was inadequately described
- not comparing like with like
- strong claims for weak data
- limited studies that examined client outcomes
- the so what factor? (ie does this make a difference/is this of any importance?)

In considering the quality of the evidence, there are other factors that need considering. A lot of the research included in this review is undertaken in the USA and hence relates to the supervision of trainees. Although supervision for trainees is a valid area of study, the majority of the evidence is from a specific population. Furthermore, the trainee experience is a time of rapid growth and development, and it could be argued that it is more difficult, therefore, to correlate any impact solely to supervision.

Another factor that needs to be considered when evaluating the relevance of the evidence is the fact that most of the research has been undertaken in the USA and the potential for the generalisability of the research evidence to the UK setting could be limited. Training curricula and training approaches in the UK and USA are different. For example, trainees in the US are often younger and in full-time training in comparison to trainees in the UK who tend to be older, because counselling tends to be a second career, and training is more often part-time. Also, the context of counselling can differ; for example, rehabilitation and school counselling do not compare readily with workplace counselling in the UK. These differences across cultural contexts will have some impact on the generalisability of results.

What is evident from these quality issues is that many of the studies in the review could have been undertaken more rigorously. There were a few studies that had a more rigorous design and these were more explicit about the methodological limitations – for example, Bambling et al. (2006). However, most of the results that are presented in this report need to be considered in light of a range of methodological weaknesses.

Detailed information about each study in the review is provided and a tentative conclusion about the quality of the study is offered. Readers may come to different conclusions by examining all the evidence provided.

4.6 Research evidence

The research evidence from the 25 studies reviewed is presented for each main impact category, with an overview of each study. This overview includes the aims of each study, the general methodological approach taken and the research tools used, along with a summary of key relevant findings of each study (see Tables 8–12). The studies are discussed in relation to each of the key impact categories and conclusions are drawn about what the research evidence claims. Information about quality of the evidence and the limitations of the studies are briefly examined. Each of the 25 studies is discussed under one category only, even though they may fit more than one category.

4.6.1 Self-awareness

4.6.1a An overview

An essential aspect of counselling and psychotherapy is the therapist's self-awareness. While different theoretical orientations use specific language to describe the phenomena, it is broadly agreed that the client makes an emotional impact on the therapist that is part of the process of the work. The therapist must be sufficiently aware of their own emotional process to be able to understand their reactions to the client's material, and to distinguish their own thoughts and feelings from those of the client in order to make sense of and use the client's communication. Hence the continuing development of self-awareness forms an integral part of counsellor training (BACP, 2003).

4.6.1b The evidence

Three studies indicated that self-awareness is enhanced through supervision. Through qualitative interviews with nine trainees, **Jimenez Andujar (2002)** explored the personal growth of supervisees. The author asked trainees how they defined and understood personal growth, what their experiences of it were and what hindered and supported these experiences within the supervisory process. She found that participants reported experiencing increased awareness through supervision, whereby some participants became more aware of personal issues. They also gained a sense of autonomy, a sense of being competent and feeling good about their abilities and were able to evaluate their performance. The conceptualisation of awareness was interrelated with the perception of participants' weaknesses and strengths in the role of counsellor, personal issues that might interfere with their performance and boundaries related to their relationships with clients and others. Participants in this study perceived individual supervision as safer than group supervision in promoting their personal growth. The study was reported well, but has all the limitations of a small-scale study with interviews conducted by the author.

Raichelson, Herron, Primavera and Ramirez (1997)

conducted a cross-sectional survey (with a qualitative item analysis of response statements) that examined parallel process in supervision, comparing psychoanalytic and non-psychoanalytic orientations. A Parallel Process Survey was designed by the authors to assess the incidence and effects of parallel process. Part of the survey gained responses about whether parallel process was perceived to exist and another part measured its value to the supervisee. There were strong correlations between the depth of perception of its existence and aspects of therapy. These included: supervisees gaining a deeper awareness of countertransference issues and subjective emotional responses to the patient ($r < .51$); supervisee ability to learn from supervisor increases ($r < .58$); the working alliance between patient and therapist is

Table 8: Research evidence for self-awareness

Authors	Aims of the study	General research approach	Sample information	Research Tools	Results that Relate to the Impact of Supervision on the Supervisee?	Overall Quality Rating
Jiménez Andújar (2002)	To explore the personal growth experiences of Masters-level supervisees within the supervisory process.	Qualitative Cross-sectional semi-structured post interviews	Supervisees 9 (3 male/6 female). 6 white American; 1 Asian-American; 1 African-American; 1 Hispanic. Mean age approximately 35. Some of the supervisors were doctoral students who had been trained in supervision.	Interviews analysed by grounded theory	Participants reported experiencing increased awareness through supervision. The conceptualisation of awareness was interrelated with the perception of participant's weaknesses and strengths in the role of counsellor, personal issues that might interfere with their performance, and boundaries related to their relationships with clients and others. The supervisees reported feeling they had more confidence and that this confidence was expressed in the way their work with their clients	Average
Raichelson, S.H., Herron, W.G., Primavera, L.H. & Ramirez, S.M. (1997)	To identify the degree to which parallel process exists in supervision and the specific effects of parallel process in psychotherapy supervision.	Mixed-methods Cross-sectional survey	Supervisors 150 (73 male/27 female). Mean age of 53.5 years. All supervisors were certified and met the requirements necessary to supervise in their respective post-doctoral training institutes. A mixture of psychoanalytic, Jungian, eclectic, rational-emotive, Cognitive Behavioural or combinations of the above. Supervisees 150 (91 male/59 female). Mean age of 48.8 years. Psychoanalytic, Jungian, eclectic, rational-emotive, cognitive behavioural or combinations of the above. The 100 psychoanalytic supervisees averaged 18.6 years conducting therapy and 10.7 years of receiving psychotherapy supervision.	Parallel Process Survey designed for the study Qualitative item analysis of form and content of statements from survey	As a result of parallel process supervisees: a) cope with negative transference feelings b) are more aware of awareness of counter-transference c) understand nonverbal behaviour d) learn from supervision e) act spontaneously and warmly in the therapeutic process.	Average
Tyron (1996)	To examine how self-other awareness, motivation and dependency change in supervision over time in practicum supervision.	Quantitative Longitudinal questionnaire: Pre, mid and post	Supervisors 2 (1 male/1 female). Psychodynamic. Supervisees 25 (8 Male/17 female). 0-2 years counselling experience. Counsellors saw on average 14 clients per week for 31 weeks.	SLQ-R (McNeil et al., 1992)	Self -other awareness: 56% had progressive scale score increases over the 3 testing Dependency autonomy: 52% had progressive scale score increases over the 3 testing Motivation: only 36% had progressive scale score increases over the 3 testing	Average

Table 9: Research evidence for skill development

Authors	Aims of the study	General research approach	Sample information	Research Tools	Results that Relate to the Impact of Supervision on the Supervisee?	Overall Quality Rating
Borders (1990)	To investigate supervisees' perceptions of development changes during their first practicum semester with regard to self-awareness, autonomy and acquisition of theory and skills.	Quantitative Longitudinal questionnaire: pre and post	Supervisees 3 (2 male/1 female). 1 "relationship-based and insight-oriented" approach. 1 Gestalt; 1 cognitive behavioural. Supervisees 44 (7 male/37 female). Mean age 36.10 years. Supervisees were all enrolled in the practicum course in a Masters counseling program (36 semester hours), having completed courses in counseling theories, skills and techniques. They had either completed or were concurrently taking additional required courses, eg group or career counselling.	Supervisee Levels Questionnaire (McNeill, Stoltenberg and Pierce, 1985)	Supervisees reported significant increases on the three dimensions of development in the model. They perceived themselves as more aware of their own motivations and dynamics, less concerned about their performance during a session and less dependent on their supervisors for directions and support. They also reported more consistent application of acquired skills and knowledge when working with clients.	Average
Carey, Williams & Wells (1988)	To test the applicability of Strong's (1968) model of training supervision by examining the relationships among supervisor expertise, attractiveness and trustworthiness, and supervision performance measures.	Quantitative Cross-sectional questionnaire: 6 weeks after starting supervision	Supervisors 17 (10 male/7 female). 7 faculty and 10 advanced doctoral. Supervisees 31 (10 male/21 female). 0-2 years' counselling experience.	Supervisor Rating Form (CRF) Heppner & Handley's (1981) – supervisee. Counselor Evaluation Rating Scale (CERS) (Myrick & Kell, 1971) – supervisor	These results are consistent with Frielander and Synder's (1983) findings that trustworthiness is most strongly associated with supervision relationship quality and suggests that this enhanced quality is reflected in enhanced trainee performance. These results also demonstrate that trustworthiness is the most salient credibility factor operating in master's practicum supervision. The results also support results of previous research reporting significant relationships between supervisor credibility and trainee performance (Dodenhoff, 1981).	Average
Dodenhoff (1981)	(a) How does interpersonal attraction between supervisor and trainee affect counselor trainee behaviour? (b) How do influencing messages influence counselor trainee behaviour? (c) What is the interaction between interpersonal attractive and influencing messages?	Quantitative Longitudinal questionnaire: pre, mid and post	Supervisors 8 of the supervisors were PhD-level faculty members experienced in supervision. 2 were doctoral-level students with past experience in supervision; 2 were Masters-level supervisors with extensive experience in supervision. Supervisees 59 (25 male/34 female).	Counselor Rating Form (Barak & LaCrosse, 1975) Rating Scale for Outcome (RSO; Storrow, 1960) Counselor Evaluation Rating Scale CERS; Myrick & Kelly, 1971).	Trainees who were attracted to their supervisors were rated as more effective by supervisors on two measures, although attraction was not related to clients' perceptions of outcomes. A direct style of supervision was related to trainee effectiveness, but on only one of three measures of the dependent variable.	Average

Authors	Aims of the study	General research approach	Sample information	Research Tools	Results that Relate to the Impact of Supervision on the Supervisee?	Overall Quality Rating
Ogren & Jonsson (2003)	To explore the attainment of psychotherapeutic skill of students before and after group supervision, and to compare supervisee supervisee ratings with supervisor ratings.	Quantitative Longitudinal questionnaire: pre and post	Supervisors Unknown – 56 supervision groups, each group comprising 3 supervisees and 1 supervisor but each supervisor had more than one group. Psychodynamic. Supervisees Sample comprised students in a basic level programme, which was part of a five-year academic training for psychologists at three universities in Sweden. Psychodynamic.	Modified Self Evaluation Scale (MSES) – A Swedish version (Olsson, 1996) of Buckley et al. (1982) Self Evaluation Scale	For the inexperienced therapist, group supervision contributes to greater skill, increasing confidence and self-esteem in the ability to handle important aspects of dynamic psychotherapy.	Good
Ogren, Jonsson & Sundin (2005)	To examine supervisees' and supervisors' views on focus and group climate in group supervision and their relationship with supervisees' attained skill.	Quantitative Longitudinal questionnaire: post only	Supervisors 35 (7 male/28 female). Mean age 53 years. Psychodynamic. All Caucasian. 2 years' training programme in supervision. Supervisees 184 (54 male/130 female). Psychodynamic. 0-2 years' experience.	Modified Self Evaluation Scale (MSES) – A Swedish version (Olsson, 1996) of Buckley et al. (1982) Self Evaluation Scale TAC – Topics and Climate Inventory (Olsson, 1996)	Attention to group process, psychodynamic processes, professional attitudes and theoretical aspects accounted for supervisees' perceptions of attained skill.	Good
Strozier, Kivlighan, & Thoreson (1993)	To examine the cognitive aspects of supervision. To assess the applicability of the intentions and reactions paradigm for supervision, as well as the usefulness of the analytic technique of sequential analysis.	Mixed-methods Questionnaires Videotaping of sessions rated	Supervisor 1 male, age 36. Psychodynamic and interpersonal. Supervisee 1 female, age 30, 2-4 years' experience. Doctoral student in counselling psychology who described orientation as interpersonal. Clients With "developmental or personality-characterological concerns".	Session Evaluation Questionnaire (Stiles & Snow, 1984) Supervisor Intentions List Hill and O'Grady's Therapist Intention List (1985). Helpfulness Rating Scale (Elliott, 1985) Supervisee Reactions System (Hill, et al, 1988)	There are meaningful sequential relationships between a supervisor's intentions and a supervisee's reactions. The sequential analyses were able to predict what led to work in supervision on the part of the supervisee. The supervisor's explore, assessment, restructure and change interventions led to supervisee's reaction of therapeutic work. Supervisee development is facilitated by focus on the relationship, support and challenge.	Average

Vallance (2004)	To explore counsellor perceptions of the impact of counselling supervision on clients.	Qualitative Qualitative questionnaires and semi-structured interviews	Supervisees 19 (1 male/18 female). White British. Person-centered bias. Mixture of newly qualified and more experienced with many years' experience.	Open-ended questionnaire and semi-structured interviews	Overall, supervision does directly and indirectly impact client work in a range of helpful and unhelpful ways. Supervision leads to: a) increased confidence, congruence, focus, freedom and safety in the client work b) professional development through increased congruence and confidence c) supervisees not being distracted by their own emotions d) supervisees monitoring their work e) development of ethical decision making f) insight into client dynamics	Average
Worthen & McNeill (1996)	To investigate good supervision events.	Qualitative Semi-structured interviews	Supervisors 8 (4 male/4 female). 1–10 years' supervisory experience. Supervisees 8 (4 male/4 female). Intermediate level to advanced level trainees. European-American ethnicity.	Interviews – supervisee	Outcomes of good supervision: ■ Strengthened confidence ■ Refined professional identity ■ Increased therapeutic perception ■ Expanded ability to conceptualise and intervene ■ Positive anticipation to re-engage in the struggle ■ Strengthened supervisory alliance	Average
Worthington (1984)	To investigate changes in perceptions of supervisees at five levels of training: after practicum 1, 2, 3, 4 and during internship. To examine supervisor impact in counselor improvement and frequency at which supervisors performed selected supervisor behaviours.	Quantitative Cross-sectional survey	Supervisors 150 Unknown but supervisors were mostly post-PhD psychologists (n=155). However, some pre-doctoral supervisors were involved (n=82), especially during the first and second practica. Supervisees 237 First semester practicum (n=31); Second semester practicum (n=67); Third semester practicum (n=53); Fourth semester practicum (n=47); Internship (n=39).	The Supervision Questionnaire (SQ) and the revised Supervision Questionnaire (SQ-R).	Perceived competences of the supervisor were related to providing acceptance and support and to teaching conceptualisation, intervention, and other skills clearly. High impact was related to personal interest (eg helping with thorny issues such as their own involvement or defensiveness). Supervision differed across settings: ■ Supervisors perceived to have high impact helped counselors get clients to accept the conceptualization more than low impact supervisors during the first three practica, but not at more experienced levels. ■ Support and encouragement were especially useful at two levels of pre-doctoral training of counseling psychologists- during first practicum and during internship.	Average

Table 10: Research evidence for self-efficacy

Authors	Aims of the study	General research approach	Sample information	Research Tools	Results that Relate to the Impact of Supervision on the Supervisee?	Overall Quality Ratings
Cashwell & Dooley (2001)	To determine what effect receiving or not receiving clinical supervision on a regular basis would have on counseling self-efficacy.	Quantitative Cross-sectional questionnaire	Supervisees 22 receiving supervision (6 male/16 female). 25–54 years old. 3 African-American; 19 Caucasian-American. 11 not receiving supervision (2 male/9 female). 23–47 years old. 2 African-American; 9 Caucasian-American.	Counseling self-estimate inventory (CPSE) Larson et al. (1992) – supervisee	Those counselors receiving clinical supervision indicated higher levels of counseling self-efficacy than those not receiving clinical supervision.	Average
Earley (2004)	To examine the relationship between clinical supervision and clinical practice by investigating how counsellors experience supervision after being licensed. This had nine research questions as part of the study. Specifically examined the type and frequency of supervision in relation to supervision anxiety and counsellor self-efficacy.	Quantitative Questionnaire Cross-sectional survey	Supervisees 210 (71 male (34%)/138 female (65%). 29–73 years old; mean age 50.38. 97% Caucasian (n=200 out of 210); 0.5% Native American (n=1 out of 210); 1.9% other (n=4).	CSSES – Counsellor self-efficacy Scale (Melchert et al, 1996) – supervisee State-Trait Anxiety Scale (-trate/STAI-state) (Spielberger, 1970) supervisee Demographics questionnaire – supervisee Survey – supervisee	An inverse correlation was found between supervision anxiety and counsellor self-efficacy. As supervision anxiety increased counsellor self-efficacy decreased. Study respondents indicated that the conditions for the development of positive counsellor self-efficacy exists in the supervision process experience by most of them, regardless of type and frequency.	Average

<p>Elstition, Patton, & Kardash (1990)</p>	<p>The development of the Supervisory Working Alliance Inventory (SWAI) to measure the relationship in counsellor supervision.</p>	<p>Quantitative Questionnaires: construction of a measure</p>	<p>Supervisors 185 (114 male/69 female/2 gender unidentified). Mean age 41.96 years. An average of 15 years of therapy experience. Psychodynamic (25%); cognitive behavioural (17%); eclectic (27%); 5 different other approaches.</p> <p>Supervisees 178 (73 male/103 female/2 gender unidentified). Mean age 29.95 years. An average of 5.70 years of therapy experience. Psychodynamic (28%); Cognitive behavioural (19%); Eclectic (16%); 5 different other approaches.</p>	<p>Supervision Working Alliance Inventory (SWAI) (Friendlander & Ward, 1984) Supervisory Styles Inventory (SSI) (Friendlander & Ward, 1984) Self-efficacy Inventory (SEI) (Friendlander & Snyder, 1983)</p>	<p>There is a significant correlation between some aspects of the Supervisory Styles Inventory, the Supervision Working Alliance Inventory and the Self-efficacy inventory: self efficacy inventory and supervision being task centred, self efficacy and rapport with the supervisor, self efficacy inventory and the attractiveness of the supervisor. The Supervision Working Alliance Inventory is a valid instrument.</p>	<p>Average</p>
<p>Jumper (1999)</p>	<p>To examine the delivery of immediate in-session feedback using the 'bug-in-the-ear' (BITE) as an instructional technique in conjunction with live supervision during the counselling practicum. The BITE technique was studied to determine its use on several key dimensions including counselling self-efficacy, trainee anxiety and skill development.</p>	<p>Quantitative Longitudinal questionnaires: pre, mid and post</p>	<p>Supervisors 3 (this included 2 doctoral students). All had a doctoral level seminar in principles and techniques of clinical supervision prior to the study. Faculty instructors provided supervision to doctoral student supervisors, as well as additional supervision to the counsellor trainees.</p> <p>Supervisees 20 beginning trainees (5 male, 15 female). Mean age 33.25 years. 19 Caucasian; 1 Native American. 0-2 years' counselling experience.</p> <p>Clients Participants counselled different types of clients, who exhibited varying degrees of motivation and resistance. Clients also had different levels of previous exposure to counselling, either at the clinic or in other settings.</p>	<p>COSE – Counseling-self estimate Inventory (Larson et al., 1992) SSTA-I: State-Trait Anxiety Scale (Spielberg, 1983). CERS – Counsellor Evaluation Rating Scale (Myrick & Kelly, 1972) TVCS – Trainee Value of Cues Scale (McClure & Virend et al, 1976). Subjective Anxiety Measure – single item measure composed for the purpose of the study. Demographic Questionnaire</p>	<p>Trainees in the BITE condition demonstrated significantly larger increases in counselling self-efficacy across the course of the study than did trainees in the no-BITE condition. Trainees in both groups experienced successful performance accomplishments during the counselling practicum and received reinforcement for these successes. However, the significant difference in BITE trainees in counselling self-efficacy provides support for the value of the immediacy of this feedback. Receiving immediate feedback during counselling session did not significantly reduce their emotional arousal over the course of the counselling practicum.</p>	<p>Average</p>

Authors	Aims of the study	General research approach	Sample information	Research Tools	Results that Relate to the Impact of Supervision on the Supervisee?	Overall Quality Ratings
Ladany, Ellis & Friedlander (1999)	To investigate whether self-efficacy and satisfaction increase as the working alliance becomes stronger.	Quantitative Longitudinal questionnaires: pre and post	Supervisees 107 (35 male/72 female). Mean age 29.91 years. 86% white; 7% African-American; 3% Latino; 2% Asian-American; 3% unknown. 0-2 years' counselling experience	Trainee personal reaction scale revised, (Holloway and Wampold, 1984) Self Efficacy Inventory - SEI (Friedlander and Snyder, 1983) Working Alliance Inventory trainee version (Bahrnick, 1990)	Changes in alliance were not predictive of changes in trainee's self-efficacy. However, improvements in the emotional bond between trainees and supervisors were associated with greater satisfaction.	Average
Lehrman-Waterman & Ladany (2001)	The development and validation of the Evaluation Process Within Supervision Inventory (EPSI).	Quantitative Questionnaires: developing and validating a measure	Supervisees 274 trainees (63 male/211 female). 223 white; 19 African-American; 11 Asian-American; 12 Hispanic; 6 bi-racial; 2 unspecified. Mean age 29.08 years. Median of 25 months of counselling experience. Supervisors 45% male/55% female. 87% white; 7% African-American; 3% Hispanic; 2% Asian-American; 1% bi-racial; 0.4% did not specify race. 85% had a doctoral or medical degree and 14% had a Masters degree.	Evaluation Process Within Supervision Inventory (EPSI).	Evidence of reliability and validity for the EPSI measure. Effective evaluation practices are: a) predictive of a stronger working alliance. b) associated with stronger perception of supervisor influence and self-efficacy. c) Predictive of greater trainee satisfaction d) Have no significant relationship with trainee training level.	Average

Ossana (1991)	To examine the impact of supervision on trainee self-efficacy and counseling ability, and testing a developmental framework.	<p>Quantitative</p> <p>Longitudinal questionnaire: pre, mid and post</p>	<p>Supervisors 47 (27 male/20 female). Mean age 36.9. Psychodynamic (average 4.6); Humanistic (average 5.5); Interpersonal orientations (average 5.5) Systems orientation (average 4.0). 40 white (85.1%) 4 black (8.5%) 2 Hispanic (4.3%) 1 Asian-American (2.1%)</p> <p>Supervisee 75 trainees (48 male/27 female). Mean age 28.76. years 64 white (85.3%); 7 black (9.3%); 4 Asian-American (5.3%). 0-2 years' counselling experience.</p> <p>Psychodynamic (average 4.22); Humanistic (average 5.30); Cognitive behavioural (average 4.64); Eclectic (average 5.62); Systems (average 4.19); Interpersonal (average 5.17).</p>	<p>Supervisee Counselor Ability Scale (CAS): a) CAS-E: supervisee self-efficacy b) CAS-A: supervisee counseling ability Personal Data Sheet Likert scale used to indicate extent to which they adhered to each of six theoretical orientations. Supervisor Counselor Ability Scale (CAS): a) CAS-A: Supervisee counseling ability b) Supervision Level Scale (SLS) (Wiley and Ray, 1986) Supervisors were asked to indicate extent to which they believed they implemented each theoretical orientation in their supervision sessions.</p>	<p>The study suggests that congruence (i.e. match of supervisor interventions and trainee characteristics) and incongruence do impact supervisee's self-efficacy and counseling ability. However, it seems that the importance of congruence to effective supervision outcome appears to vary according to the developmental dimensions (i.e beginning or advanced) that are being evaluated as well as to who (ie supervisor or trainee) perceives the congruence and when (i.e onset or conclusion of supervision) the congruence is perceived.</p>	Average
---------------	--	---	--	--	---	---------

Table 11: Research evidence for transfer of understanding from supervision to therapy

Authors	Aims of the study	General research approach	Sample information	Research tools	Results that relate to the impact of supervision on the supervisee	Overall Quality Rating
Couchon & Bernard (1984)	To investigate whether the timing of supervision makes a difference to the outcome with counsellor and client.	Mixed-methods Questionnaires Video tapes	Supervisors 7 (4 male/3 female). Age range 25–42 years. Supervisees 21 Masters students (4 male/17 female).	Mainly quantitative but supervision tapes were used to compile future oriented strategies list. Counseling Evaluation Inventory. Video recording – supervisee and supervisor. Client satisfaction with counseling.	Supervision sessions conducted just prior to counseling (4 hours) appeared to have been more focused and seemed to be more of a 'planning session' for upcoming counseling. Follow through from supervision to counseling was greatest at this time. Supervision sessions held the day before counseling appeared to be more content-orientated, characterised by a focus on conceptual material taught by the supervisor. Because of the quantity of material in these supervision sessions, counselor follow through from supervision to counseling was lower at this time.	Average
Guest & Beutler (1988)	To investigate the relationship between changes in the theoretical orientation and values of psychotherapy trainees and the orientation and values of their supervisors.	Mixed-methods Questionnaires Cross-sectional survey	Supervisors 11 psychodynamic and cognitive behavioural. Supervisees 16 trainees (8 full-time pre-doctoral interns and 8 advanced pre-internship graduate students within context of a doctoral programme).	Theoretical Orientation Questionnaire (Sundland, 1977a, 1977b) Value Survey (Rokeach, 1973, 1979) Locus of Control (Rotter, 1966) Personality (Eysenck & Eysenck, 1969)	The results showed: (a) that beginning trainees value support and technical direction (b) that with experience trainees come to value supervisors who hold complex, dynamic views of change (c) that advanced trainees place increasing value on the assessment of personal issues and relationships that may affect the psychotherapy process	Average
Milne, Pilkington, Gracie & James (2003)	To assess the effectiveness of CBT supervision in terms of its observed impact on a supervisee and her patient.	Qualitative Single case study design Video recording (with quantitative content analysis)	Supervisor 1 male, age 33 years. Cognitive behavioural therapy. 2 years' experience as a supervisor. Supervisee 1 female, mid 40s.GP with no prior counselling experience.	Video recording – supervisee, supervisor and client. Cross-sectional design with 20 tapes from the same dyad being used. Qualitative collection of data with some quantitative analysis of data.	The study provides limited evidence that supervision in CBT can be effective, as measured by appropriate changes in therapy consequent upon supervision. There is considerable thematic transference of an appropriate kind from supervision to therapy.	Average

<p>Patton & Kivlighan (1997)</p>	<p>To examine the extent to which the trainee's perception of the supervisory working alliance is related to outcomes of the supervisory process: (a) the client's perception of the working alliance in counselling and (b) the trainee's adherence to the prescribed counselling approach.</p>	<p>Mixed-methods Questionnaires Observation Videotapes</p>	<p>Supervisor 25 (7 male/18 female). 27-41 years (mean age 32.30). All 25 were European Americans. An average of 103 hours of supervisory experience. Interpersonal supervision.</p> <p>Supervisees 75 (22 male/53 female). Mean age 27.71 years.</p> <p>Students enrolled in graduate-level prepracticum university course. Some had been in paraprofessional counselling positions; for most, however, this was their first formal counselling experience. 0-2 years' experience Psychodynamic.</p>	<p>Supervisor Working Alliance Inventory (Efstation et al., 1990) Working Alliance Inventory (Horvath & Greenberg, 1989). Observation and videotaping of supervisee.</p>	<p>The supervisory working alliance has a differential impact on the types of learning that occur in supervision. The supervisory working alliance not only indexes the trainee's comfort in the supervisory relationship but also is related to the trainee's performance in counselling.</p>	<p>Average</p>
<p>Steinheiber, J., Patterson, V., Cliffe, K. & Legouillon, M. (1984)</p>	<p>To investigate the amount of supervision and the congruence of therapeutic orientation between the therapist, the supervisor and the therapy and patient change.</p>	<p>Quantitative Cross-sectional questionnaires post only</p>	<p>Supervisees 51 trainees in a number of mental health disciplines.</p> <p>Clients 237 (83 male/54 female). 13.1% psychosis; 31.2% neurosis; 8.4 borderline personalities; 31% other personality disorder; 12.7 transient situational disturbances; 3.4% other diagnoses.</p> <p>Mean number of sessions 31.1, (SD 28.9) over 8.4 months (SD=6.1).</p>	<p>Global Assessment of Functioning (Endicott, Spitzer, Fleiss and Cohen, 1976) Patient problem and diagnosis</p>	<p>Congruent supervision was more frequent with patients seen one or more times weekly. Prescription of medication was related to non-congruent supervision. Amount of supervision was not related to patient change.</p>	<p>Average</p>

Table 12: Research evidence for client outcome

Authors	Aims of the study	General research approach	Sample information	Research Tools	Results that Relate to the Impact of Supervision on the Supervisee?	Overall Quality Rating
Bambling, King, Raue, Schweitzer & Lambert (2006)	To examine the impact of clinical supervision on client working alliance and symptom reduction in the brief treatment of major depression. It also examined any differences in client outcomes between process- vs. skill-focused supervision. The study examined 5 related hypotheses to these aims.	<p>Quantitative</p> <p>Randomised Control Trial</p>	<p>Supervisors 40 (9 male/31 female). Average age 49.9 years (SD=9.4, range 34–97 years).</p> <p>Supervisees 127 (31 male/96 female). Counselling experience mean 8.8 years (SD=5.8, range=2–25). Minimum therapists' requirements were graduate qualifications in mental health and 1 year graduate experience of providing counseling services to clients.</p> <p>Clients 127 (40 male/87 female). Inclusion criterion: primary diagnosis of major depression using MINI Inventory and BDI. Comorbidity along Axis I and Axis 11 of the DSM-IV was tolerated as long as major depression was the primary diagnosis. Existing medication users were accepted into this study because pre-existing antidepressant use has not been shown to either enhance or detract from positive treatment outcome in CBT. Antidepressant use was the basis for exclusion if started at the time of intake because it might confound the interpretation of gains made in treatment. 24 clients (18.9%) were current users of antidepressant medication – 103 clients (81.1%) were not using medication for depression. 66 clients (52%) had a diagnosis of major depression only. Primary comorbidity: agoraphobia (n=4, 3.1%); social phobias (n=5, 3.9%); obsessive-compulsive disorder (n=33, 26%) and panic disorder (n=14, 11%). Other primary Axis II Diagnosis revealed a range of personality disorders.</p>	<p>The Social Skills Inventory (SSI) (Crowley, 2000) – supervisee</p> <p>Problem Solving Adherence Scale – supervisee</p> <p>Supervision Focus Adherence Scale (SFAS) – supervisor</p> <p>Beck Depression Inventory – client</p> <p>Treatment Evaluation Scale – client</p> <p>Working Alliance Inventory – client</p> <p>Intake assessment – client</p>	<p>The results of this study provide qualified support that supervision that focuses on working alliance can influence client perception of alliance and enhance treatment outcome in the brief psychotherapeutic treatment of depression. At this stage, there is evidence that both the skill and process supervision used in this study were able to effectively achieve these ends.</p>	Excellent

strengthened ($r < .52$); patients examine resistances that were previously avoided in therapy ($r < .47$); and the supervisory relationship is strengthened and communication improves ($r < .60$) (p.44). In summary, supervisees could be described as gaining in self-awareness through the use of the parallel process.

This is a cross-sectional study with sampling problems – there was less than 25% response to invitation to participate. No causal effect can be concluded. The sample included relatively highly experienced therapists and supervisors. More work needs to be done on the survey instrument to improve validity and reliability. The survey provides only self-report information and does not relate to specific observable cases.

Tyron (1996) used a longitudinal design to investigate supervisee development over the course of a year of training practice, using the 30-item SLQ-R (McNeill et al, 1992) to assess supervisee self-other awareness (12 items), motivation (8 items) and dependency-autonomy (10 items). 56% of the participants in the study had increased in self-other awareness scores over a period of time, indicating that supervision can positively impact upon the self-awareness of the supervisee. 52% had progressive score increases on the dependency-autonomy scale but only 13% had progressive scores for motivation. Limitations of the study are various. It was a small sample (25 supervisees), data was collected from five different supervisees each year over five years, they were all from the same training course and the questionnaires were self-report only. An independent observer may have weighted them differently. Their ratings may have been influenced by their expectations of development over the year.

4.6.1c A summary of the evidence

Although all three studies conclude that self-awareness is increased through supervision, the limitations of the studies reduce confidence in the findings. The development of self-awareness is an essential aspect of counsellor training and would be expected for all trainees. There are numerous stimuli for self-development: personal therapy, transmission of knowledge, relationships with fellow students, relationships with clients and life events, as well as supervision. None of these studies are controlled in such a way that could result in the claim that only supervision was responsible for the increase in self-awareness. Only Tyron (1996) used a longitudinal design that measured change over time, but this was a small sample and there were many design limitations. The overall conclusion is that self-awareness develops as a result of the counselling training and some of that development may be attributed to supervision.

4.6.2 Skill development

4.6.2a An overview

One of the prime purposes of supervision is to facilitate the development of techniques and skills that are an integral part of the therapeutic process. This is particularly pertinent when the supervisees are trainees. Nine studies provided some evidence of skill development as a result of supervision.

4.6.2b The evidence

Borders' (1990) quantitative investigation (with a pre- and post-questionnaire design) into supervisees' perceptions of their development during their first semester of practice found that supervisees reported a more consistent application of acquired skills and knowledge when working with clients as a result of supervision. The Supervisee Levels Questionnaire (SLQ; McNeill et al, 1985) was completed by students in the

second and sixteenth weeks of one semester. There were three supervision groups, each with a different supervisor. A preliminary analysis found no significant differences between the developmental level of supervisees in the three groups. Supervisees in all three supervision groups reported significant increases in self-awareness, dependency/autonomy and theory/skills acquisition. These results indicated that supervisees experienced significant increases through supervision on three developmental dimensions: they perceived themselves as more aware of their own motivations and dynamics, they were less concerned about their performance during a session and were less dependent on their supervisors for direction and support. Limitations of the study included the relatively small number of participants from one counselling programme and the reliance on one self-report measure of developmental change.

Carey, Williams & Wells (1988) conducted a quantitative study (where a questionnaire was administered to supervisees six weeks after starting supervision) to test 'the applicability of Strong's (1986) model to supervision by examining the relationships among supervisor expertness, attractiveness and trustworthiness and supervision performance measures' (p132). They collected data from 31 Masters-level trainees after they had completed six sessions of supervision and found that enhanced trainee performance is influenced by the level of trustworthiness the supervisee experiences the supervisor as having. The Counsellor Evaluation Rating Scale (CERS) was developed to measure practicum trainee performance based upon supervisor ratings. Statistically significant (p.05) relationships were found between supervisor trustworthiness and all six CERS subscales (p.135). The author notes that these findings are consistent with those of Friedlander and Synder (1983), and that supervisor trustworthiness is of particular importance at the beginning of Masters practicum because supervision at this time helps trainees manage anxiety and promotes trainee risk-taking, along with experimentation. There are numerous limitations to this study, including the lack of independent measures of supervisor qualities and supervisee performance. Measures were only used at one point in time and conclusions can only tentatively be drawn about trainees at the beginning of their practicum. There was a small sample size of just 31 supervisor/supervisee pairs.

In a quantitative longitudinal study, **Dodenhoff (1981)** explored supervision as social influence process, hypothesising that counsellor attraction to the supervisor would lead to greater supervisor influence and subsequently to improved client outcome (Freitas, 2002: p.354). Trainee attraction to the supervisor was found to positively correlate with the supervisor's evaluation of client outcome and the supervisor's evaluation of supervisee skill (p.52). Clients' ratings of therapy outcome bore no relationship to the supervisor attractiveness or supervisor directiveness. Both Freitas (2002) and Efstation (1990) critique the study, noting a range of limitations. These include: a lack of information on psychometric measures and data, supervisors judging client outcomes, the supervisors' judgements of therapeutic outcomes differed from clients' perceptions of outcomes, and the training level of supervisors was varied, hence the difficulty in drawing meaningful inferences.

In a quantitative study (pre- and post-questionnaire design) on group supervision of therapists conducting therapy with individuals, **Ogren and Jonsson (2003)** explored the attainment of psychotherapeutic skill of students before and after group supervision, and compared supervisee ratings with supervisor ratings. The study results suggested that supervision contributes to greater skill, with the likelihood that this skill development enhances a supervisee's ability

to manage key aspects of the psychotherapeutic process. Supervisees rated their psychotherapeutic skill before and after supervision with a higher rating afterwards. This outcome was valid for the three factors that were explored (ability to contain and deal with emotionally loaded therapeutic issues, mastery of working alliance and psychodynamic understanding). Furthermore, the differences in skill level were considerable in 'ability to contain and deal with emotionally loaded issues and in psychodynamic understanding, amounting to about 1 and 0.5 Standard Deviation of the mean raw scores respectively. For the inexperienced therapist, group supervision therefore contributes to greater skill, and increased confidence and self-esteem in the ability to handle important aspects of dynamic psychotherapy (p. 54).

Ogren, Jonsson and Sundin (2005) also undertook a quantitative study (questionnaire design – post only) of group supervision and found that supervisees' perceptions of attained skill were a result of attention to group process, psychodynamic processes, professional attitudes and theoretical aspects. Hierarchical regression analysis showed that differences in supervisor style were related to supervisees' experience of attained skills. Supervisees' experience of the degree to which focus in supervision was on psychodynamic process was positively related to perceived attained skill, whereas focus on theoretical aspects was negatively related. A theoretical focus appeared to have a negative effect on the supervisee's experience of attained psychotherapeutic skill in handling emotional issues, as they experienced less ability in this area (p.382). The results of the study also showed that supervisor ratings of focus in supervision were significantly higher than the supervisees' ratings on all dimensions (p.386). There are various limitations to the study. The Topics and Climate (TAC) questionnaire was developed for this and tested on a small sample of 37 participants, insufficient to be sure of its reliability. One of the subscales of the TAC had poor psychometric properties. Also, Buckley's Self Evaluation Scale (Buckley et al, 1982) was used, translated into Swedish with some items removed, which may have affected its reliability. However, both of the above studies indicate that individual and group supervision does improve counselling skills in a number of different ways.

Using a single-subject and mixed-method design, **Strozier, Kivlighan and Thoreson (1993)** examined the intentions of supervisors and the reactions of supervisees to these. The findings suggested that there were meaningful sequential relationships between a supervisor's intentions and a supervisee's reactions. Intentions and reactions were coded through the sequential analysis of video-recorded sessions. The supervisor's explore, assessment, restructure and change interventions led to a reaction on the part of the supervisee that was observable in her therapeutic work. The supervisee development seemed to be facilitated by focus on the relationship, support and challenge. There are numerous limitations to this study, not least that the supervisor was also the researcher and second author of the paper. As a single case study, the results are just an indicator of possible links that need to be explored in larger scale studies. It only explored cognitive approaches and relates to the experiences of an advanced trainee and experienced supervisor. There are also limitations to sequential analysis, as only the immediate impact of a response can be captured and only categorical data can be used.

A qualitative study by **Vallance (2004)** explored counsellors' perceptions of the impact of counselling supervision on work with clients. From 13 qualitative questionnaires completed by counsellors and six semi-structured interviews with different counsellors, she found that emotional support from supervision benefits counsellors and directly impacts on client

work. This is because it ensures that counsellors are not distracted by their own emotions and prevents contamination of the client processes, indicating an increase of therapeutic competence. She also identified that supervision leads to increased confidence, congruence, focus, freedom and safety in the client work, and increased confidence had a positive impact on the professional development of the supervisee. The study has all the limitations inherent in small-scale qualitative studies. There is potential for researcher bias as the counsellors were known to the researcher and had similar training and shared values with her. The impact on clients is only as the counsellors report it.

In a qualitative study that explored the experience of 'good' supervision from the perspective of supervisees, **Worthen and McNeill (1996)** found that good supervision experience strengthened the confidence and refined the professional identity of the supervisee. Eight trainees participated in tape-recorded interviews in which they described a recent good supervision experience. Four distinct supervision phases of a good supervision experience for advanced supervisees were identified. These were existential baseline, setting the stage, good supervision experience and outcomes of good supervision. For this review, the phase of 'outcomes of good supervision' is most relevant. The authors report that in each of the individual cases there were many expressions regarding the effects of good supervision, but six outcome themes were consistently identified from the supervisees in the study (p.31). These outcomes were: 1) strengthened confidence, 2) refined professional identity, 3) increased therapeutic perception, 4) expanded ability to conceptualise and intervene, 5) positive anticipation to re-engage in the struggle and 6) strengthened supervisory alliance. In the article, narratives from some of the trainees interviewed are presented under each theme. Although a small sample size, all trainees report a range of positive outcomes of supervision. The limitations of the study include the small sample size. As the authors note, even though the study was a phenomenological investigation to describe and uncover the structures of personal meaning, it is not generalisable nor can it predict all good experiences of supervision and therefore may not be inclusive of all important events or elements of what may be considered good supervision (p.33). The study only focused on the perspective of the supervisee and the authors point out that a different version of good supervision might emerge from supervisors' perspectives. Finally, the study doesn't link experience of good supervision with whether supervisees actually become more effective in facilitating better therapy outcomes in clients – this is a key issue for many of the studies included in this review.

In a quantitative cross-sectional survey, **Worthington (1984)** examined changes in the perception of supervision by supervisees at five different levels of training from first practical experience through to post-qualification. Measures included satisfaction with supervision, supervisor competence and supervisor impact on the counsellor. This study involved a large and diverse sample (237 supervisor/supervisee pairs) and generated a considerable amount of data and statistical tests. Some supervisor behaviours (eg encouragement, feedback about facilitative and non-facilitative behaviours, and appropriate confrontation) seemed to promote satisfied and effective counsellors albeit with varied training and experience. Conceptualisation and intervention skills appear to have been learned at all experience levels. There are numerous limitations to this study, not least its complexity. The lack of homogeneity of subjects revealed numerous differences across settings and theoretical orientation. Also, Ellis and Ladany (1997) have criticised this study, saying that the problems are so severe that they obviate making inferences.

4.6.2c A summary of the evidence

Skill development as a descriptor for this category of study is broadly defined to be an enhanced capacity to interact with a client in a helpful way. The articles reported explored diverse aspects of skill-development measures or those recorded in different ways, including self-report by the trainee, supervisor report, interviews, observations and subjective opinion. Studies included a single-subject case study, and group and individual supervision. The measurement of skill development was not the specific focus of many of the studies, but was reported as part of other findings. None of the studies are rated as of more than average in quality. The qualitative studies were more rigorously conducted, but report on small studies. They also risk researcher bias, as in some studies participants were known to the interviewer. It is reasonable to expect that counselling and psychotherapeutic skills will develop through supervision, as part of its role and function is to be educative (Inskipp and Proctor, 2001). Research participants report this to be the case, but a rigorous study that links skill development with supervision is yet to be designed, undertaken and reported.

4.6.3 Self-efficacy

4.6.3a An overview

Self-efficacy can be described as the therapist's belief in their own capacity for effective practice. Counsellors' self-efficacy beliefs are "the primary causal determinant of effective counselling action" (Larson and Daniels, 1998: 180), and are, therefore, an important component of supervision. Efficacy in dealing with the environment (or with a client) does not entail certain behaviours or simply knowing what to do. It involves a capability that encompasses social, cognitive, conceptual and behavioural skills that can be mobilised in the service of the client or the agency in which the work is carried out. Counsellors are required to harness their knowledge, experience and self-awareness to make decisions related to circumstances that can be many and various and can change. They need to have faith in their own capacity (or self-efficacy) to function effectively as a therapist.

Self-efficacy is an implicit feature of many studies of supervision, given that developing this in the therapist is a fundamental aim of supervision. In this review, seven studies included supervision and self-efficacy as all or part of the research enquiry.

4.6.3b The evidence

In a quantitative study (post-questionnaire only), **Cashwell and Dooley (2001)** posed the hypothesis that counsellors receiving supervision would have a higher level of self-efficacy than counsellors not engaging in supervision. The Counselling Self-Estimate Inventory (COSE) (Larson et al, 1992) was used at one point in time with a total of 33 participants – 11, without supervision and 22 with supervision. Their hypothesis was confirmed. A statistically significant difference was found between counsellors receiving clinical supervision and those not receiving it at $p=.024$. Those counsellors receiving clinical supervision indicated higher levels of counselling self-efficacy that those not receiving supervision. The authors only state small sample size as a limitation of the study. However, other limitations are evident. The sample is biased towards those having supervision – 22 included in the study are having supervision and only 11 were not having supervision. Also, there is no examination of any variables, nor is self-efficacy monitored over time to measure change. Further, it is still not clear which supervisees from the doctoral programme and from the agency setting are in the 'not having supervision' group or the 'having supervision' group. The authors state

that some students had an administrative supervisor but this difference was not distinguished in any way. The 22 counsellors having supervision were seeking their professional licences, whereas the group not receiving supervision were not, but reasons for this difference were not given. In conclusion, the comparisons are weak, with limited sample size, and factors related to the counselling setting not being reported.

In a quantitative study (using questionnaires and a survey), **Earley (2004)** examined the relationship between clinical supervision and clinical practice by investigating how counsellors experience supervision after being licensed. The study examined the type and frequency of supervision in relation to anxiety about supervision and counsellor self-efficacy. An inverse correlation was found between anxiety about supervision and counsellor self-efficacy; as supervision anxiety increased, counsellor self-efficacy decreased. Most of the respondents indicated that conditions for the development of positive counsellor self-efficacy exist in the supervision process, regardless of type and frequency. The authors note a range of study limitations relating to sample, study design and instruments used. The survey was a convenience sample and self-report data was collected. It was an ex-post facto (retrospective) correlational study that does not establish a causal relationship among the variables. The Counsellor Self-Efficacy Scale (CSES) (Melchert et al, 1996) scale is a task-related inventory that does not investigate non-task factors that may be associated with performance, such as the personal attributes of the counsellor. The State-Trait Anxiety Scale (STAI) (Spielberger, 1970) reports participants' general self-assessment of expected anxiety in supervision rather than a direct measure of immediate, experienced supervision anxiety. As a result, it does not capture the range of anxiety levels that may be experienced by a counsellor in supervision over a series of sessions. Some of the questions in the Demographics Questionnaire were forced choice questions rather than scaled or open-ended questions. These questions did not allow respondents to specify the degree of conviction they held for the choice selected or to identify reasons for the choices they made. The study was limited to one-off data collection.

Efstation, Patton and Kardash (1990) designed their quantitative study to construct a measure of the supervisor's and the trainee's perception of the supervisory relationship. They invited 614 counsellor-training directors to participate in the study and received complete and usable responses from 33% of their target participants. As part of the battery of instruments used, trainees completed a supervision outcome measure, the Self-Efficacy Inventory (SEI; Friedlander and Snyder, 1983). The self-efficacy scale refers to trainees' feelings about their capacity and confidence to perform assessments, engage in therapy and manage crises. In this study the trainees had a considerable amount of experience (mean 5.7 years). A significant correlation was found between supervision being task centred, there existing a rapport with the supervisor and the attractiveness of the supervisor with self-efficacy. Limitations of the study include the fact that data was collected between one supervisor and one supervisee at one point in time. The supervisees were considerably experienced and nothing can be concluded about trainees at other points of their developmental process. The therapeutic alliance develops over time and little can be concluded from data collected just once. The sample included supervisors of various theoretical orientations who may stress different dimensions of supervision according to the trainees' level of experience and the requirements of the theoretical model.

In a quantitative study (using a pre, mid and post questionnaire), **Jumper (1999)** examined the delivery of

immediate in-session feedback using the 'bug-in-the-ear' (BITE) as an instructional technique in conjunction with live supervision during the counselling practicum. The results of this study suggest that the immediacy of the feedback via the BITE process increases supervisees' self-efficacy. Trainees receiving immediate in-session feedback reported that cues that were reinforcing or encouraging were most helpful. The authors suggest that this has a clear link to the development of self-efficacy, because "according to self-efficacy theory, successful performance accomplishments have the most powerful influences on an individual's perceptions of self-efficacy" (Bandura, 1977). The study sample size was somewhat small – with 10 trainees receiving the BITE condition and 10 trainees not receiving it (but still receiving supervision). These results are, therefore, not generalisable to all trainees.

Ladany, Ellis and Friedlander (1999) designed a longitudinal quantitative study to test Bordin's (1983) proposition that changes in trainees' perceptions of the supervisory alliance over the course of supervision would predict supervisory outcomes. The results showed significant gains over time in self-efficacy, but changes in the alliance, taken together or considered separately, did not predict changes in self-efficacy. However, improvements in the emotional bond between the trainees and supervisors were associated with greater satisfaction. The authors discuss many limitations of the study, including the fact that self-efficacy may increase regardless of the relationship with the supervisor, as training, peer influence and other forms of learning may have an impact. This could be said of any study of supervision – it does not happen in a vacuum. The authors also note the threats to validity from ex-post facto designs. Predictor variables are not manipulated and participants are not randomly assigned to conditions that threaten internal validity. The results can only be generalised to trainees with similar demographic characteristics to those in the study. The supervisory process was only measured from the trainee's perspective and that of the supervisor or others may have yielded different results.

Leherman-Waterman and Ladany (2001) conducted a quantitative study to develop and validate the self-reported measure to assess the supervisees' experience of goal setting and feedback aspects of evaluation within the supervisory process. As part of this process, they collected self-efficacy data using the Self-Efficacy Inventory (SEI) (Friedlander and Snyder, 1983). Supervisees were asked to complete the battery of questionnaires they were sent based on recall of a supervisory relationship. The participants were 74 counselling trainees with a range of counselling experience (Median 25 months). Results relevant to this review are that there were significant correlations between supervisees' self-efficacy and the ways in which supervisors set goals and gave feedback. There were numerous limitations to this study. The construct 'evaluation' may have been too narrowly defined. Reliance on a single measure, single method approach, at one point in time, limits what can be deduced from the findings. There was a 35% return rate and the problem of who self-selected to complete the survey. Participants were required to recall the supervisory relationship and the effect of evaluation from memory. The final limitation is the ex-post facto nature of the design that precludes causal inferences.

In a quantitative study (using questionnaires at pre, mid and post times), **Ossana (1991)** tested developmental models of supervision. One of the three research questions of the study investigated whether greater congruence (ie match of supervisory intentions with supervisee developmental characteristics) was predictive of supervision outcome (defined in this study as supervisee self-efficacy and counselling ability). The results indicated that the more

supervisors perceived congruence between their beginning interventions (eg structure instruction) and beginning trainee characteristics (eg anxiety, insecurity) early in the supervisory process, the more trainees were perceived as improving with regard to self-efficacy and counselling ability (p.172). Also, the more trainees' perceived congruence between beginning supervisory interventions and their beginning developmental characteristics at the conclusion of supervision, the more their self-efficacy and perceived counselling ability improved (p.173). Both supervisors' and trainees' ratings of congruence were, therefore, associated with supervision outcome, providing some support for the importance of congruence to effective supervision (p. 175).

4.6.3c A summary of the evidence

Self-efficacy was not the primary focus of research in most of these studies. One sought to validate a measure (Lehrman & Ladany, 2001) and another sought primarily to construct a measure (Efstation et al., 1990). The study by Jumper (1999) sought to investigate a technique, and Ossana (1991) set out to test the development model. While the study by Cashwell and Dolley (2001) addresses directly the question of whether supervision made a difference to self-efficacy, regrettably the research design, choice of subjects and data collection were such that little can be reliably concluded from the study. Two studies used the Self-Efficacy Inventory (Friedland and Snyder, 1983), otherwise all studies used different measures of self-efficacy.

There is ample scope to use reliable instruments repeatedly in order to establish whether supervision impacts on self-efficacy reliably in different contexts with varied supervisees in terms of experience level and other variables. Such work is yet to be done. There is a lack of longitudinal studies measuring change over time and of well-designed controlled trials. Too often, data is collated retrospectively, requiring supervisees to recall a supervisory relationship. In summary, the studies reported provide weak evidence that supervision enhances self-efficacy in counsellors.

4.6.4 Transfer of understanding from supervision to therapy

4.6.4a An overview

Supervision has the potential to have an impact on many behaviours of the therapist, as well as outcome for the client. One aspect of such influence could be the way in which the therapist internalises the values, beliefs, skills and behaviours of the supervisor, which may be radically different in supervision of trainee therapists than supervision with experienced practitioners. Such transfer of understanding may also be enhanced depending on the timing of supervision. The influence of the supervisor could potentially be observed in the subsequent sessions with clients. Three studies have explored the way in which theoretical orientation is absorbed by the therapist. Two other studies look at the impact of supervision as a result of its timing with respect to client sessions.

4.6.4b The evidence

In a mixed-method study (using questionnaires and video tape recordings of the supervisee and supervisor), **Couchon and Bernard (1984)** investigated whether the timing of supervision made any difference to the outcome of supervision. They found that the timing of supervision did impact on the content of supervision. For example, the content of supervision conducted immediately prior to counselling (four hours, same day) was more focused on planning for upcoming counselling sessions. In comparison, the content of supervision the day

before a counselling session was more characterised by a focus on conceptual material taught by the supervisor, and also yielded a high volume of information, most of which was generated by the supervisor. Overall, issues discussed in supervision immediately prior to counselling seemed to transfer more to the counselling session than supervision the day before counselling. Efstation (1990) describes some of the limitations of this study: some counsellors saw the same clients under different conditions; some supervisors listened to audio tapes whereas others viewed video tapes; there were no controls for Type 1 error (the error of rejecting a null hypothesis when it is actually true); there was minimal description of measures and no psychometric data was presented.

In a mixed-method study (using questionnaires and a cross-sectional survey), **Guest and Beutler (1988)** examined the relationship between changes in the theoretical orientation and values of psychotherapy trainees, and the orientation and values of their supervisors. They collected questionnaire data from trainee therapists and supervisors. Multiple regression analyses were used to evaluate the contribution of supervisors' theoretical orientations and personal values to changes in the trainees' theoretical orientation. The results suggested that beginners value support and technical direction and that, with experience, trainees come to value supervisors who hold complex, dynamic views of change. Advanced trainees place increasing value on the assessment of personal issues and relationships that may affect the psychotherapy process. The implications of this research point towards supervisors applying different supervisory models at different levels of training. It was also found that supervisors' orientations exerted a reliable influence on trainees' theoretical orientations three to five years following the conclusion of their training experience. The authors note limitations to the study to be that the statistical power of the study was limited by its small sample size (16 trainees) and by the large number of variables examined. It wasn't entirely clear what the total number of supervisors in the sample was and the assertions made need to be treated with some caution, especially as sub-groups of the sample of 16 trainees included eight beginning and eight experienced trainees.

In a single case study design, **Milne, Pilkington, Gracie and James (2003)** assessed the effectiveness of CBT supervision and found that appropriate thematic transference from supervision to therapy occurred. The frequencies of themes observed in supervision and transferred to therapy were counted. Across 10 observed tapes of supervision and therapy, between four and nine themes were found to transfer from each supervision session to the subsequent therapy session (mean=5.6; standard deviation =1.67). The most frequently observed form of transfer occurred in relation to the informing/educating themes, which occurred during all observed sessions. Only 7.4% of themes (on two observed occasions) were transferred inappropriately. It was also found that some of the transfer observed reflects the supervisee's increasing awareness of how various CBT methodologies can promote change. However, as pointed out by Milne et al. (ibid: 200), this does "not support the notion that a complete 1:1 transfer of actions from supervision to therapy necessarily makes for good therapy". The authors note a range of study limitations, including the fact that the work focused on only one of the supervisees from the dyad. Thus, the study is not representative of supervisees. Also, the raters of all the 20 tapes were not themselves qualified in either CBT or supervision, which may have contributed to biases or errors in recording. For note, however, the first author – an experienced supervisor – did make some attempts to check for reliability of the data analyses with another researcher.

In a mixed-method study (using questionnaires, observations and video tapes of supervisees), **Patton and Kivlighan (1997)** compared the trainee's perception of the supervisory alliance to the client's perception of the counselling alliance and to the trainee's adherence to the model of counselling promoted through supervision. This study involved the use of volunteer clients, but as they were seen for four sessions, this study has been included in the review. Had it been fewer sessions and hence less like a 'real' client setting, the study would not have met the inclusion criteria. Significant relationships were found as predicted and trainees appear to develop knowledge about building and maintaining relationships that is used in the relationship with clients. They also demonstrated that the supervisory working alliance has a differential impact on the types of learning that occur in supervision. The significant linear change coefficient indicated that client WAI – Working Alliance Inventory (Efstation et al, 1990) – scores increased each session and thus clients saw the working alliance as increasing linearly across the four counselling sessions. The variance components confirmed the t-test results, indicating that the supervisory working alliance accounted for 27% of the variance in 'between session' ratings of trainee adherence to general psychodynamic interviewing style. Efstation et al (1990) describe some of the limitations of this study, which include questions about statistical procedures. Also, as sessions were observed through a one-way mirror, and supervisor comments were offered during the sessions, clients' perceptions of therapy may be influenced by the participation of the supervisor, as well as the interventions of the therapist, and therefore the conduct of the therapy could have been influenced by the supervisor's advice/interventions.

In a longitudinal quantitative study, **Steinheber, Patterson, Cliffe and Legouillon (1984)** reported that congruence of theoretical orientation between supervisor and trainee counsellor was related to patient change as measured by the Global Assessment Scale (GAS, Endicott et al., 1976). The GAS scale is used by the therapist to assess and record the severity of the patient's problems. Means scored for patients in the congruent group changed from 62.5 (SD=13.7) to 68.6 (SD=12.6), compared with the smaller change of the non-congruent group from 59.4 (SD=13.3) to 61.1 (SD=13.9). For the purpose of statistical analysis, supervisees were allocated to three different supervision groups depending on the amount of supervision they received during the period of the study (low, medium and high). Congruence of theoretical orientation between supervisor and trainee counsellor was found to be related significantly to GAS change in the low supervision group but not in medium or high groups. The study concluded that congruence of supervision, but not the amount of supervision, was related significantly to GAS change in the cases studied (p.1356). It also found that the amount of supervision bore a positive relationship to patient attendance. Efstation (1990) describes considerable limitations of this study: trainees varied considerably in the number of patients they saw; patients varied in diagnosis; the amount of supervision varied depending on what seemed to be required (eg more supervision for more difficult clients); there was an overly simplistic categorisation of supervisory congruence and patient wellbeing outcome was assessed by trainees.

4.6.4c A summary of the evidence

It is not surprising that supervisees, particularly those in training, are influenced by their supervisors and supervision is likely to reinforce the learning from any training undertaken, providing that there is congruence between the theoretical training model, the theoretical model used in supervision and what is required or expected in practice. These studies

demonstrate the influence the supervisor has on the counsellor and the therapeutic work. The quality of the evidence is weak for numerous reasons: small samples, lack of homogeneity of samples and strategies, contrived therapy sessions and single-point data collection. Possibly the most robust study, Milne et al (2003) used a single case design that mapped the influence of supervision over time. However, there were methodological problems that need to be noted. Overall, the evidence supports the assertion that learning in supervision is transferred to practice, but not reliably.

4.6.5 Client outcome

4.6.5a An overview

Ultimately, counselling and psychotherapy are conducted for the benefit of the client. The most powerful demonstration of the efficacy of supervision is manifest in changes in the client that can be attributed to the way in which the therapist's work is enhanced through supervision. Several studies attempt to explore this relationship but are reported under other categories. Only one is described here.

4.6.5b The evidence

The most recent study to be included in this review was conducted by **Bambling, King, Raue, Schweitzer and Lambert (2006)**. This is the only Randomised Controlled Trial (RCT) in the review. Its aim was to investigate the impact of supervision on client outcome, examining the impact of clinical supervision on the working alliance between therapist and client, and symptom reduction in the brief treatment of major depression. As part of the study, the authors also compared process- vs. skill-focused supervision to identify whether there were any differences between the type of supervision and client outcome. The study involved 127 clients diagnosed with major depression who were randomly assigned to 127 supervised or unsupervised therapists. It used a nested study design with multiple intervals of measurement and had the following five hypotheses:

1. 'Depressed clients who receive supervised therapy will demonstrate higher working alliance scores compared with control groups receiving unsupervised therapy.' This hypothesis was supported.
2. 'Depressed clients receiving supervised therapy will demonstrate a greater reduction of Beck Depression Inventory (BDI) scores compared with a control group receiving the same type of therapy from unsupervised therapists.' This hypothesis was supported.
3. 'Clients receiving supervised therapy will evaluate their treatment more positively compared with a control group of clients who receive the same type of therapy but from unsupervised therapists.' This hypothesis was supported.
4. 'Clients receiving supervised therapy will be more likely to complete treatment compared to the control group of clients who receive the same type of therapy but from unsupervised therapists.' This hypothesis was supported.
5. 'The supervision process approach that focuses on creating awareness of interpersonal processes to manage the working alliance in therapy will demonstrate different clients' working alliance scores and reduction of BDI scores at treatment endpoint than skill-focused supervision procedures focused on behavioural strategies to manage the working alliance'. This hypothesis was not supported.

Supervision manuals were developed for the study in both process-focused and working-alliance, skill-focused mode.

The results of this study provided qualified support that supervision focused on the working alliance can influence client perception of the alliance and enhance treatment outcome in the brief psychotherapeutic treatment of depression. Clients receiving supervised therapy achieved a significantly greater reduction in Beck Depression Inventory (BDI) scores than those receiving unsupervised therapy. Also clients who were treated by supervised therapists were more satisfied than those treated by unsupervised therapists. There was evidence that both skill and process supervision used in this study were able to effectively achieve these ends. The results suggest, therefore, that relational and insight techniques are neither superior nor inferior to cognitive and behavioural techniques.

The intention was to have 65 discrete supervisors supervising 65 therapists each having their own client. A pre-study power calculation was conducted but the anticipated power was not achieved, as the required number of clients were not recruited. Concealment of allocation was not clear. In practice, Bambling supervised 13 therapists himself, with equal numbers in each supervision mode. Volunteer supervisors were trained in one or other mode of supervision. Therapists were also volunteers, as were clients who responded to media advertising and local medical health networks.

Limitations of the study were reported to be as follows. There was a pre-treatment session for therapists conducted by a group of supervisors (unspecified) during which supervisors instructed therapists on early alliance management principles and therapists discussed characteristics of the case (they had access to client assessment information and histories) with an emphasis on applying alliance management principles. This may have had an impact on the therapists together with therapist allegiance effects (unspecified), and the pre-treatment alliance training session might have provided a superior starting point for working alliances. Secondly, although 103 clients completed the therapy, we cannot be certain that Type II errors (failing to reject a null hypothesis when the alternative hypothesis is true) did not occur, as the sample was smaller than that intended in the design of the study. Thirdly, the study could not detect subtle differences over time in working alliances and symptom changes. Hence it was impossible to detect whether any tensions in therapy were resolved at any particular point through supervision. Fourthly, the study lost 18.9% of patients to follow up, and it is not clear whether the study conducted an 'intention to treat' analysis. Finally, the researcher was a supervisor in the study, which may have affected results in some way that cannot be clarified.

One major implication is that, given that supervision seemed to have a significant effect on the therapeutic alliance, training in alliance-focused supervision could be valuable for future supervisors. The other major implication is that trials of supervision are possible and can show positive effects.

4.6.5c A summary of the evidence

This is the only study in the review that has been ranked highly for quality. There is evidence that the authors reviewed the relevant literature well and that they tried to design a study that would overcome many of the methodological difficulties inherent in other studies of supervision. Some of the difficulties they encountered are common to psychotherapy research: a recruitment of subjects, dropouts and Type II errors. Nonetheless, the study presents the best evidence yet that supervision has a beneficial effect on the supervisor, the client and the outcome of therapy.

Section 5: Conclusions and implications for research and practice

The focus for this review – ‘the impact of supervision on supervisees, their clients and their practice’ – was chosen specifically to draw together and critically appraise relevant research that could provide an evidence base for the efficacy of clinical supervision. It has succeeded in the task set, but the outcome does not provide the robust evidence that would ideally be required to make bold statements about the efficacy of supervision practice. Based on existing evidence, supervision would be unlikely to be included in NICE (National Institute of Clinical Evidence) Guidelines, as it does not have enough of ‘best quality evidence’ to support the need for it to enhance and support the therapeutic process. Nonetheless, the review will equip trainers, policy makers, managers and supervisors to quote a range of diverse studies in support of their arguments for supervision service provision. This review purposely excluded studies reporting on satisfaction with supervision. There would be many such studies, but the decision was made that such subjective reports would not be considered reliable evidence. Given the dearth of robust studies supporting the efficacy of supervision, perhaps satisfaction should have been included as a category, because at the very least this report would have been able to make a statement about how much supervisees value the supervisory process.

The studies in this review do, however, indicate that supervision has a positive impact on the supervisee, whereby they grow and develop. It is further evident that supervision has some impact on key developmental areas, such as skills, self-awareness and self-efficacy. The findings indicate that change can be measured and that there are some reliable instruments available to measure it. The strength of these findings needs to be considered in relation to the quality and methodological rigour of the studies. Furthermore, most studies involve trainees, whose development may be influenced by other factors in their training environment.

The usefulness of these findings to practice is mixed. Supervision does seem to offer opportunities for supervisees to improve practice and gain in confidence, which raises the likelihood that client outcome is improved as an indirect result of supervision. However, the link to improved outcome for clients is tentative and, apart from one study (Bambling et al, 2006), there are no studies in this review that offer substantial evidence to support improvement in client outcomes. Furthermore, the majority of studies examined impact over relatively short periods of time; the longer term impact of supervision is not reflected in these studies.

The major contribution this review can make for practice is to the future strategy for supervision research. It highlights the many deficiencies in existing research and paves the way for future studies that can address valuable questions, using methods that take account of the mistakes of others. There are numerous potential supervision research questions that could enhance our knowledge and evidence base, but suggestions made here relate particularly to the focus of this review – ‘the impact of supervision on the therapist, their practice and their clients.’

There is inevitably scope for more randomised controlled trials of supervision. The study by Bambling et al. (2006) compared therapists with and without supervision working with clients presenting with major depression. This study could be replicated using therapists working with any other

client group. However, there are potential ethical difficulties with such studies in the professional culture of the UK that expects all practitioners to receive regular supervision. Such difficulties could possibly be addressed with studies across professional cultures; for example, comparing similar groups in the UK and the USA, where there is not an expectation of lifelong supervision. It might also be possible to compare therapists from different professions, such as counselling and psychiatry, working with similar clients, as psychiatrists are less likely to routinely receive supervision. Another way of addressing the ethical problem might be to compare supervised and unsupervised cases within the caseload of particular therapists. The therapist would be receiving supervision but not on a sample of cases. It is certainly a priority to find a subject group of therapists that will enable a well-organised randomised controlled trial to be conducted. This might provide an indication of the impact of supervision, given a British sample.

It was surprising that we found no studies that relate supervision to therapist wellbeing. Orliinsky & Ronnestad (2005) have reported on aspects of therapist wellbeing in their international survey of the development of psychotherapists, but little can be deduced about the impact of supervision. International comparative studies could be designed to compare therapists with and without supervision with dependent variables such as job satisfaction, wellbeing, personal development, professional development, creativity as therapists, capacity to reflect on client work and ethical decision-making to name but a few. Such research would also provide more information and insight into ways in which experienced practitioners benefit from supervision.

There is plenty of scope to design projects that investigate the impact of particular types of supervisory interventions on the performance of counsellors or therapists as measured by client outcome. For example, Worthen & Lambert (2007) discuss the impact of providing feedback on the progress of clients as regularly measured by the OQ 45 through supervision. Interventions such as viewing video recordings, listening to tape recordings, using process notes in supervision or using client monitoring data could all be compared to routine supervision. Using client monitoring data such as OQ 45 (ibid) or Clinical Outcomes in Routine Evaluation (CORE) (see www.coreims.co.uk) studies could be designed to compare the outcome of clients who are taken or not taken to supervision. The same supervisor would be involved with the same therapist reducing the confounding factors if length of therapy, client-presenting problem and severity were matched. While in the UK supervision is a career-long requirement, there is no imperative to take every client to supervision.

From another perspective, supervisor variables and the impact on therapists and clients could be explored. Variables might include whether the supervisor has undergone supervision training or their congruence with the therapeutic orientation and training of the therapist. The roles, tasks and functions that the supervisor performs (Inskipp & Proctor, 1995) as well as the supervisory style could be explored with the counsellor’s capacity for reflection, competence and client outcome.

Given the methodological complexity of such quantitative research, there is also scope for well-designed and rigorously conducted qualitative research, including in-depth case studies. Tape recordings of therapy sessions and subsequent supervision sessions can reveal much about ways in which supervision influences practice. Interviews or

focus groups with therapists that explore the subtleties of their experience of supervision either in the present or over time may provide new insights that support the efficacy of supervision. As shown in this review, although qualitative studies cannot claim generalisability, they can supply important detail of supervisory practice and supervisee experience. Any kind

of research on supervision for experienced practitioners and across the career lifespan would be very welcome. A key outcome of this review is recognition of the need to formulate clearly defined research questions, to engage in methodological plurality including triangulation and last, but not least, to include the voice of the client.

Section 6: Research articles

6.1 Abstracts of articles/studies included in the review

Study	Bamblings, 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
Item type	Journal, Article
Journal	<i>Psychotherapy Research</i>
Date of publication	2006
Volume	16
Issue	3
Pages	317–331
Abstract	Supervision of psychotherapists and counsellors, especially in the early years of practice, is widely accepted as being important for professional development and to ensure optimal client outcomes. Although the process of clinical supervision has been extensively studied, less is known about the impact of supervision on psychotherapy practice and client symptom outcome. This study evaluated the impact of clinical supervision on client working alliance and symptom reduction in the brief treatment of major depression. The authors randomly assigned 127 clients with a diagnosis of major depression to 127 supervised or unsupervised therapists to receive eight sessions of problem-solving treatment. Supervised therapists were randomly assigned to either alliance skill- or alliance process-focused supervision and received eight supervision sessions. Before beginning treatment, therapists received one supervision session for brief training in the working alliance supervision approach and in specific characteristics of each case. Standard measures of therapeutic alliance and symptom change were used as dependent variables. The results showed a significant effect for both supervision conditions on working alliance from the first session of therapy, symptom reduction, and treatment retention and evaluation but no effect differences between supervision conditions. It was not feasible to separate the effects of supervision from the single pre-treatment session and it is possible that allegiance effects might have inflated results. The scientific and clinical relevance of these findings is discussed.
Study type	Quantitative: Randomised Control Trial (RCT)
Definition of supervision	Yes, but tentatively. During supervision, a supervisor and therapist may systematically examine case-specific treatment and process issues as a method of enhancing both therapist awareness and skills necessary to manage the complexities of client work.
How were participants recruited?	<i>Supervisors:</i> volunteers recruited through private setting. <i>Therapists:</i> therapists recruited through media and journal advertisement. <i>Clients:</i> volunteers recruited through local mental health networks and media advertising.
Dropouts	24 clients (18.9%) did not complete the full course of therapy, leaving 103 completed client cases.
Incentives	<i>Supervisors:</i> given free training accredited by the University of Queensland Psychiatry Department that could be used for Continuing Education (CE) points for other professional purposes. <i>Therapists:</i> given free accredited training in PST in return for participating in the study. <i>Clients:</i> treatment for free.
Limitations	1) It is not possible to separate the supervision effect from any pre-treatment session and therapist allegiance effects. 2) Although providing a moderate sample size, total power was insufficient to ensure that Type II errors could not occur. 3) The principal researcher in this study undertook a significant amount of the supervision across both conditions – it remains possible that researcher provided supervision might have influenced results in some unknown way, reducing the detectable effect between approaches.
Our quality rating	Excellent.

Study	Borders, L. D., Williams, K. S., & Wells, M. (1990). Developmental changes during supervisees' first practicum
Item type	Journal, Article
Journal	<i>Clinical Supervisor</i>
Date of publication	1990
Volume	8
Issue	2
Pages	157–167
Abstract	Investigated short-term changes along three dimensions postulated by C. Stoltenberg (see PA, Vol 65:6282): self-awareness, dependency/autonomy, and theory/skills acquisition. 44 supervisees (aged 22 to 58 yrs) completed the Supervisee Levels Questionnaire of B. W. McNeill et al (see PA, Vol 73:10234) at the beginning and end of their 1st practicum. An analysis revealed supervisees reported statistically significant gains on all three developmental dimensions. Results provide additional support for the developmental perspective on supervision, but indicated a need to clarify the definition of developmental level.
Study type	Quantitative: pre and post questionnaire, supervisor and supervisee.
Definition of supervision	None provided.
How were participants recruited?	Participants were drawn from 46 students enrolled in the practicum course in a Masters counselling program at a medium-sized university in the American Midwest. All students volunteered to participate but two did not complete the post-test.
Dropouts	Two did not complete the post-test.
Incentives	None reported.
Limitations	Limitations included the relatively small number of participants from one counseling program and the reliance on one self-report measure of developmental change.
Our quality rating	Average.

Study	Carey, J. C. (1988). Relationships between dimensions of supervisors' influence and counselor trainees' performance
Item type	Journal, Article
Journal	<i>Counselor Education & Supervision</i>
Date of publication	1988
Volume	28
Issue	2
Pages	130–39
Abstract	Examined relationships among supervisor expertness, attractiveness and trustworthiness and counselor trainee performance. Subjects were 17 supervisors and 31 Masters degree-level trainees. Results showed that practicum supervisors' expertness, attractiveness and trustworthiness were related to supervision outcome measures. Trustworthiness had the strongest associations.
Study type	Quantitative: questionnaire six weeks after supervisee started supervision, supervisee and supervisor.
Definition of supervision	None provided.
How were participants recruited?	Six training courses were invited to participate.
Dropouts	None stated.
Incentives	None stated.
Limitations	None stated, and very little information was provided on numerous aspects of the study.
Our quality rating	Average.

Study	Cashwell, T. H., & Dooley, K. (2001). The impact of supervision on counselor self-efficacy
Item type	Journal, Article
Journal	<i>Clinical Supervisor</i>
Date of publication	2001
Volume	20
Issue	1
Pages	39–47
Abstract	Examined the effect of the receipt of regular clinical supervision on counseling self-efficacy. Thirty-three counsellors (aged 23–55 yrs) in community and university settings, of whom 22 were receiving clinical supervision, completed the Counseling Self-Estimate Inventory (L. M. Larson et al, 1992). Results show that those receiving regular clinical supervision reported a higher level of counselling self-efficacy than those who did not.
Study type	Quantitative: questionnaire at one point in time, supervisee.
Factors evaluated in the study	The impact of receiving clinical supervision versus receiving no clinical supervision on the counselling self-efficacy of practising counsellors. Counselor self-efficacy: micro-skills, counselling process, dealing with difficult client behaviours, cultural competence and values.
Definition of supervision	None provided.
How were participants recruited?	Fifty-four packs inviting participants were distributed to those in an agency, whereby contact was initially made with the person in charge of staff development for permission. (No rationale/info as to why this setting.) The packs contained a letter explaining the study, informed consent form, demographic info sheet and Counseling Self-Estimate Inventory). Packs were also sent to doctoral students interested in participating.
Dropouts	Unclear.
Incentives	Not stated.
Limitations	The authors just state a small sample. Other limitations include: the sample is biased towards those having supervision – 22 included in study were having supervision, whereas only 11 were not. Also, there is no examination of any variables, nor is self-efficacy monitored over time to see whether any changes exist for the sample group. Also, it is not clear whether those in doctoral programme/from agency setting are in the 'not having supervision' group or the 'having supervision' group. Furthermore, it says some students had an administrative supervisor, but the authors don't distinguish this difference in anyway. The 22 counsellors receiving supervision were seeking their professional licenses whereas the control group (who were not receiving supervision) were not. The reason for this is not given, but it could result in a possible difference in the way the sample as a whole viewed themselves and their self-efficacy. Data was collected at one point only. There was a limited sample size. Also, factors related to the setting of the counselling are not discussed in detail.
Our quality rating	Average.

Study	Couchon, W. D., & Bernard J. M. (1984). Effects of timing and supervision on supervisor and counselor performance
Item type	Journal, Article
Journal	<i>Clinical Supervisor</i>
Date of publication	1984
Volume	2
Pages	3–20
Abstract	This study examined the effect of timing of supervision on the nature of supervision and subsequent counselling. Three levels of treatment were investigated for differential effects on (1) supervisor and counsellor performance in supervision, (2) counsellor performance in counselling, (3) client and counsellor satisfaction with counselling, and (4) counsellor satisfaction with supervision. Fifty-five supervisor-counselor-subject triads received treatment. All supervision and counselling sessions were audio-taped and rated for critical dependent variables. It was hypothesized that supervision occurring within four hours before counselling (T1) would be significantly different both from supervision occurring the day before counselling (T2) and from supervision occurring more than two days before counselling (T3). Results indicated that T2 yielded a high volume of information, most of which was generated by the supervisor. T1 yielded less information because of a decrease in activity on the part of the supervisor. However, follow-through from supervision to counselling as measured by counsellor behaviour was greatest for T1. There was no significant difference among treatments regarding client satisfaction or counsellor satisfaction.
Type of study	Mixed methods: questionnaire and video tapes of supervisee and supervisor.
Definition of supervision	Yes, but tentatively. During supervision, a supervisor and therapist may systematically examine case-specific treatment and process issues as a method of enhancing both therapist awareness and skills necessary to manage the complexities of client work.
How were participants recruited?	All participants were involved as either staff or clients at the Counseling and Guidance where the research was undertaken.
Dropouts	There were 63 subjects: supervisor-counsellor-client triads. Eight were unable to be included because of scheduling difficulties.
Incentives	Not stated.
Limitations	Little detail was given about instruments used for data collected and the involvement of clients. Statistical limitations were also noted.
Our quality rating	Average.

Item	Dodenhoff, J. T. (1981). Interpersonal-attraction and direct-indirect supervisor influence as predictors of counselor trainee effectiveness
Item type	Journal, Article
Journal	<i>Journal of Counseling Psychology</i>
Date of publication	1981
Volume	28
Issue	1
Pages	47–52
Type of study	Quantitative: questionnaire (pre, mid and post), supervisee and supervisor.
Definition of supervision	“Face-to-face contact with another licensed counsellor or group of counsellors for the purpose of clarifying, evaluating and improving case management, clinical skill or conceptualisation of clinically related issues commonly considered relevant to the practice of counselling.” (Magnuson, Norem & Wilcoxon, 2000)
How were participants recruited?	Trainee participants (students in a MA in Counselling) enrolled in a counseling practicum during the fall semester of 1977. Clients were using the Arizona State University Counselor Training Center to receive free personal and/or vocational counselling.
Dropouts	Three students dropped out of their placement during the semester because of scheduling difficulties. Twelve students were eliminated from the study due to incomplete data.
Incentives	Free counselling was provided by Arizona State University Counselor Training Center.
Limitations	None stated.
Our quality rating	Average.

Study	Earley, T. M. (2004). Counselor supervision post-licensure: relationship of clinical supervision type and frequency to supervision anxiety and counselor self-efficacy
Item type	PhD
Date of publication	2004
Abstract	<p>Increasingly, mental health clinicians, including counsellors, clinical social workers and psychologists, are being expected by consumers, insurers and other professionals to demonstrate competence in the practice of counseling. Although this expectation seems self-evident, it has traditionally been met only through the initial credentialing process along with periodic documentation of ongoing practice and continuing education. Some credentialing bodies, however, have concluded that these criteria are necessary but insufficient criteria to assure continued competence after licensing. Clinical supervision has become the fourth requirement considered to ensure competent performance by licensed counsellors. This study examined the relationship between clinical supervision and clinical practice by investigating how counsellors experienced supervision after being licensed. Nine hypotheses about the relationships among type and frequency of clinical supervision, supervision anxiety and counsellor self-efficacy were tested. A survey consisting of a counselor demographics questionnaire, the Counselor Self-Efficacy Scale, and the State-Trait Anxiety Inventory was completed by 210 active licensed clinical professional counsellors in Maine. The findings and their relevance to the practice of counseling, counselor supervision and counsellor re-licensing requirements are discussed. Recommendations for further research are presented.</p>
Type of study	Mixed methods: questionnaire and survey, supervisee.
Definition of supervision	None stated.
How were participants recruited?	<p>All 649 licensed clinical professional counsellors listed in the Maine Board of Counselor Licensure's public record actively practicing in Maine.</p> <p>A survey was sent to all on this list.</p>
Dropouts	None stated.
Incentives	None stated.
Limitations	<p>These were related to sample, design and instruments used. Survey was a convenience sample. Data collected was self-reported. It was an ex-post facto, correlation study that cannot establish a causal relationship among the variables.</p> <p>The CSES scale was a task-related inventory and, therefore, does not investigate non-task factors that may be associated with performance, such as the personal attributes of the counselor.</p> <p>The STAI reports participants' general self-assessment of expected anxiety in supervision rather than a direct measure of immediate, experienced supervision anxiety. As a result, it does not capture the range of anxiety levels that may be experienced by a counsellor in supervision over a series of supervision sessions.</p> <p>The Demographics Q'aire – questions numbered 11–15 were forced choice questions rather than scaled or open-ended. These did not allow respondents to specify the degree of conviction they held for the choice selected or to identify reasons for the choices they made. There was a one data collection point.</p>
Our quality rating	Average.

Item	Efstation, J. F., Patton, M. J., & Kardash, C. M. (1990). Measuring the working alliance in counselor supervision
Item type	Journal, Article
Journal	<i>Journal of Counseling Psychology</i>
Date of publication	1990
Volume	37
Issue	3
Pages	322–329
Notes	We developed the Supervisory Working Alliance Inventory (SWAI) to measure the relationship in counselor supervision. SWAI was based conceptually on the work of R. R. Greenson (1967), H. B. Pepinsky and M. J. Patton (1971), and others. Three supervisor factors (Client Focus, Rapport and Identification) and two trainee factors (Rapport and Client Focus) were extracted by factor analysis. The scores on the SWAI were found to possess adequate scale reliability, and evidence of convergent and divergent validity for the SWAI was established by examining its relation to selected scales from the Supervisory Styles Inventory (M. L. Friedlander and L. G. Ward; see PA, Vol 72:5618). Trainee scores on the Rapport and Client Focus scales of the SWAI were significant predictors of scores on the Self-Efficacy Inventory (Friedlander and J. Snyder; see PA, Vol 71:4972). Implications for counselor training are discussed in the context of additional research on the psychometric properties of the SWAI.
Type of study	Quantitative: questionnaire – concurrently, supervisee and supervisor.
Definition of supervision	“Supervisor acts purposefully to influence trainees through their use of technical knowledge and skill and in which trainees act willingly to display their acquisition of knowledge and skill” (Pepipstay and Patton, 1990).
How were participants recruited?	The 614 training directors of internship programs listed in the Directory: Internship Programs in Professional Psychology (Association of Psychology Internship Centres, 1985–1986) and directors of graduate training programs in counselling and clinical psychology listed in Graduate Study in psychology and Associated Fields (American Psychological Association, 1986) were contacted by mail and asked to participate anonymously in an investigation of the relationship between a supervisor and a trainee in counselling and therapy supervision.
Dropouts	Not stated.
Incentives	None.
Limitations	Limitations of the study include the fact that data was collected between one supervisor and one supervisee at one point in time. The supervisees were considerably experienced and nothing can be concluded about trainees at other points of their developmental process. The therapeutic alliance develops over time and little can be concluded from data collected just once. The sample included supervisors of various theoretical orientations who may stress different dimensions of supervision according to the trainees’ level of experience and the requirements of the theoretical model.
Our quality rating	Average.

Study	Guest, P. D., & Beutler, L. E. (1988). Impact of psychotherapy supervision on therapist orientation and values
Item type	Journal, Article
Journal	<i>Journal of Consulting & Clinical Psychology</i>
Date of publication	1988
Volume	56
Issue	5
Pages	653–58
Abstract	Reviews literature on the impact of psychotherapy supervision on therapist effectiveness. It notes that beginning trainees tend to value and benefit from supervisor support while later trainees gravitate toward technical guidance, and that this cycle may be repeated with continuing experience. It presents preliminary data to suggest that early supervisory relationship has long-term effects on how psychotherapists view and approach treatment.
Type of study	Mixed-methods: questionnaire and survey, supervisee and supervisor.
Definition of supervision	None provided.
How were participants recruited?	<p>An initial training sample was drawn from clinical psychology graduate students, from a one-year psychology training program within a medical school's Department of Psychiatry. Supervised training cases were assigned through the Psychiatry Department outpatient clinic, which operated a sliding-scale, fee-for-service basis.</p> <p>Initial data were collected on all clinical psychology trainees who completed one year of supervised training during three successive training years. The sample used for analyses was based on 16 trainees. One half of this trainee sample (n=8) comprised full-time pre-doctoral interns.</p> <p>The remaining (n=8) were advanced pre-internship graduate students pursuing a paid, half-time, one-year outpatient psychotherapy training experience within the context of a doctoral program in clinical psychology. Supervisors included all training program faculty who had provided adult psychotherapy supervision to the three training cohorts (p655).</p>
Dropouts	None stated.
Incentives	Free supervision for taking part in the study.
Limitations	The authors note that the statistical power of the study was limited by its small sample size (16 trainees) and by the large number of variables examined. It was not clear what the total number of supervisors in the sample was and the assertions made need to be treated with some caution, especially as sub-groups of the sample of 16 trainees included eight beginning trainees and eight experienced trainees.
Our quality rating	Average.

Study	Jiménez Andújar, D. J. (2002). Experiences of personal growth of Masters-level supervisees within the supervisory process
Item type	PhD Dissertation, University of New Orleans
Date of publication	2002
	<p>The purpose of this study was to explore the personal growth of experiences of Masters-level supervisees within the supervisory process. Most past research in counselling supervision has been theoretical in nature and has focused on developmental models of supervision. There is a limited quantity of literature in the field that would help supervisors and counsellor educators better understand the subjective experiences that constitute supervisees' self-perceived personal growth. Understanding these experiences could enhance the quality of supervision activities provided to Masters degree students within counselling programs.</p> <p>In this study, grounded theory procedures were utilised, to develop a theoretical conceptualisation of personal growth experiences of Masters-level supervisees within the supervisory process. Two individual interviews were conducted with each of nine students from three CACREP-accredited counselling programs in the state of Louisiana. Interview data were taped and later transcribed to facilitate the extrapolation of themes.</p> <p>Findings from this study indicated that supervisees perceived their experiences of personal growth through a process called self-perception of changes. Supervisees outlined characteristics they observed in themselves throughout the supervisory process. To experience personal growth in supervision, supervisees regarded supervision as a necessarily safe environment. Supervisees identified the supervisor's personal and professional characteristics as important in promoting personal growth within the supervisory process. Supervisees valued individual supervision over group supervision, as they felt safer, more trusting and thus closer to the supervisor in individual supervision than in group supervision.</p>
Type of study	Qualitative: post interviews, supervisees.
Definition of supervision	None provided.
How were participants recruited?	Two individual interviews were conducted with each of nine Masters-level students from three CACREP-accredited counselling programs in the state of Louisiana.
Dropouts	None stated.
Incentives	None stated.
Limitations	<p>Self-report of experience cannot be generalised or measured.</p> <p>Small sample.</p> <p>No information about the therapy.</p> <p>No information about supervisors.</p> <p>Questions were woolly.</p> <p>The answers were purely subjective.</p>
Our quality rating	Average.

Study	Jumper, S. A. (1999). Immediate feedback using the 'bug-in-the-ear' in counselor training: implications for counseling self-efficacy, trainee anxiety and skill development
Item type	PhD Dissertation, University of North Dakota
Date of publication	1999
Abstract	<p>This investigation examines the delivery of immediate, in-session feedback using the 'bug-in-the-ear' (BITE) as an instructional technique in conjunction with live supervision during the counseling practicum. The study was conducted to explore an effective means of supervisor intervention, which did not disrupt the counseling session. Few empirical investigations have been conducted in this area, and previous studies on this instructional aid used in models of live supervision were largely narrative in design. Counselling self-efficacy, trainee anxiety and counselling performance were examined for 20 graduate student counsellor trainees enrolled in the department of counselling at a Northern Plains university. Ten participants received immediate feedback via the BITE in conjunction with a live supervision model of training during the first half of 10 practicum sessions conducted at a community counseling clinic. Ten participants serving as controls received live supervision without the BITE feedback during their 10 sessions. Results indicated that participants who received immediate feedback via the BITE demonstrated significantly greater increases in counseling self-efficacy throughout the course of the investigation than did the control group participants. Changes in participant anxiety levels did not differ significantly between groups. BITE or no-BITE feedback anxiety levels did not differ significantly between groups. BITE or no-BITE feedback condition, changes in counseling self-efficacy and changes in anxiety level combined to account for significant portions of the variance in participants' scores on two measures of counselling performance. Participants reported no adverse effects due to the immediate feedback, although problems with the physical equipment were noted. A series of exploratory analyses based on previous BITE investigations were also conducted. Attempts to theoretically explain the benefits of incorporating immediate feedback in live supervision using Bandura's (1977) self-efficacy theory are presented. Implications for the training of graduate students in the counselling practicum and suggestions for future research in this area are discussed. ((c) 1999 APA/PsyclINFO, all rights reserved).</p>
Type of study	Quantitative (pre, mid and post).
Definition of supervision	None provided.
How were participants recruited?	A total of 20 graduate students registered for the Masters counseling practicum course at the University of North Dakota Department of Counseling. It doesn't offer any further comments on recruitment or why this sample was used.
Dropouts	None stated.
Incentives	None stated.
Limitations	The study sample size was somewhat small – with 10 trainees receiving the BITE condition and 10 trainees not receiving the BITE condition (but still receiving supervision). These results are, therefore, not generalisable to all trainees.
Our quality rating	Average.

Study	Ladany, N., Ellis, M. V., & Friedlander, M. L. (1999). The supervisory working alliance, trainee self-efficacy and satisfaction
Item type	Journal, Article
Journal	<i>Journal of Counseling and Development</i>
Date of publication	1999
Volume	77
Issue	4
Pages	447–455
Abstract	Theoretically, when the supervisory working alliance is strong, the trainee and supervisor share a strong emotional bond and agree on the goals and tasks of supervision. The theory tested was Bordin's (1983) proposition – that changes in counsellor trainees' perceptions of the supervisory alliance over the course of supervision would predict supervisory outcomes. A national sample of beginning practicum- to intern-level trainees were assessed at the beginning and end of an academic semester. Contrary to predictions, changes in the alliance were not predictive of variations in trainees' self-efficacy. However, improvements in the emotional bond between the trainees and supervisors were associated with greater satisfaction.
Type of study	Longitudinal: quantitative
How were participants recruited?	Personal contacts in university clinical, counselling psychology programmes in several USA states. Potential volunteers were asked to participate in a study concerning the process of supervision.
Dropouts	There were 44 packets that were not usable, resulting in 107 complete sets.
Incentives	None stated.
Limitations	The authors also note the threats to validity from ex-post facto designs. Predictor variables were not manipulated and participants were not randomly assigned to conditions that threaten internal validity. The results can only be generalised to trainees with similar demographic characteristics to those in the study. The supervisory process was only measured from the trainee's perspective, and that of the supervisor or others may have yielded different results.
Our quality rating	Average.

Study	Lehrman-Waterman, D., & Ladany, N. (2001). Development and validation of the evaluation process within supervision inventory
Item type	Journal, Article
Journal	<i>Journal of Counseling Psychology Special Issue</i>
Date of publication	2001
Volume	48
Issue	2
Pages	168–177
Abstract	The purpose of this study was to develop the Evaluation Process Within Supervision Inventory (EPSI), a measure that examines evaluation practices in clinical supervision. Results based on 274 trainees with diverse training experiences indicated that the measure yields two theoretically consistent factors: goal setting (eg goals that are specific; feasible in regard to capacity, opportunity and resources; and measurable) and feedback (eg which is systematic, timely, clear and balanced between positive and negative statements). Evidence of validity was based on theoretically consistent relationships between more effective goal-setting and feedback practices and (a) a stronger working alliance, (b) enhanced trainee perception of supervisor influence on self-efficacy and (c) increased trainee satisfaction with supervision.
Type of study	Quantitative: questionnaires, supervisee and supervisor.
How were participants recruited?	Potential participants were sent a questionnaire packet consisting of an incentive postcard, a cover letter, a copy of the EPSI, the demographic form and the SSQ. Half received the SE-I and half the WAI-T in addition to the other items. All participants completed EPSI, a demographic form, the SSQ and one additional scale.
Dropouts	None stated.
Incentives	A postcard informing participants that three of the returned postcards will be randomly selected to receive \$25. Postcards with names and addresses could be returned separately from the research questionnaires.
Limitations	There were numerous limitations to this study including the following: the construct 'evaluation' may have been too narrowly defined; reliance on a single-measure, single-method approach, at one point in time limits what can be deduced from the findings. There was a 35% return rate and the problem of who self-selected to complete the survey. The participants were acquired to recall the supervisory relationship and the effect of evaluation from memory; and, finally, the ex-post facto nature of the design that precludes causal inferences.
Our quality rating	Average.

Study	Milne, D. L., Pilkington, J., Gracie, J., & James, I. (2003). Transferring skills from supervision to therapy: a qualitative and quantitative N=1 analysis
Item type	Journal, Article
Journal	<i>Behavioural and Cognitive Psychotherapy</i>
Date of publication	2003
Volume	31
Issue	2
Pages	193–202
Abstract	The importance of clinical supervision is increasingly recognised, as is the need to demonstrate that the work we do in the NHS is effective. However, observational analyses of supervision or evidence bearing out the effectiveness of supervision are rare and narrow in focus. To contribute to the evidence base, the present N=1 content and outcome evaluation describes and assesses the effectiveness of cognitive behavioural therapy supervision, in terms of its observed impacts on a supervisee and her patient. The supervisee was a trainee therapist participating in a diploma course in cognitive therapy. The study utilised a qualitative and quantitative content analysis methodology, based on the intensive coding of a series of 10 longitudinal, video-recorded supervision sessions, linked to the subsequent 10 therapy sessions. Based on this method, 14 supervisory themes were extracted, which served to describe the change methods employed in the supervision. The predicted transfer (generalisation) of those themes from supervision to therapy was observed to occur to a surprisingly marked extent, indicating that the supervision was effective. Implications are drawn for developing supervision and related research.
Type of study	Qualitative (with quantitative content analysis): supervisee and supervisor.
Definition of supervision	None provided.
How were participants recruited?	One supervisee from a training programme who was in supervision with a peer. The other supervisee in the pair was away sick for one week, but no further explanation was given as to why this pair were considered in the first place. The data used had been collected for a previous study involving tape recordings. No information was provided about the sampling process.
Dropouts	None stated.
Incentives	None stated.
Limitations	The authors note a range of study limitations, including the fact that the work focused on only one of the supervisees from the dyad. Also, the raters of the 20 tapes were not themselves qualified in either CBT or supervision, which may have contributed to biases or errors in recording. For note, however, the first author – an experienced supervisor – does some inter-rater reliability work with data analysis.
Our quality rating	Average.

Study	Ogren, M. L., Jonsson, C. O., & Sundin, E.C. (2005). Group supervision in psychotherapy: the relationship between focus, group climate and perceived attained skill
Item type	Journal, Article
Journal	<i>Journal of Clinical Psychology</i>
Date of publication	2005
Volume	61
Issue	4
Pages	373–88
Abstract	This study examined supervisees' and supervisors' views on focus and group climate in group supervision and their relationship with supervisees' attained skills. After supervision, supervisees completed a revised version of Buckley's measure of psychotherapeutic skill. Supervisees' and supervisors' experience of focus and climate in the supervision was measured with a questionnaire, Topics and Climate (TAC). The results showed that supervisors' ratings of the extent to which different foci were used in the supervision were significantly higher compared to supervisees' ratings. Hierarchical regression analysis showed that differences in supervisor style were related to supervisees' experience of attained skill. Supervisees' experience of the degree to which focus in supervision was on psychodynamic processes was positively related to perceived attained skill, whereas focus on theoretical aspects was negatively related. These findings underline the importance of research work with topics and group processes in group supervision. AD – Stockholm University, Sweden.
Type of study	Quantitative: questionnaire (post only), supervisee.
Definition of supervision	None provided.
How were participants recruited?	All students in the last two years of a five-year professional psychology program.
Dropouts	Yes.
Incentives	None stated.
Limitations	There are various limitations to the study. The Topics and Climate (TAC) questionnaire was developed for the study and tested on a small sample of 37 participants, which is insufficient to be sure of its reliability. One of the subscales of the TAC had poor psychometric properties. Also, Buckley's Self-Evaluation Scale (Buckley et al, 1982) was used and translated into Swedish with some items removed, which may have affected its reliability. However, both of the above two studies indicate that individual and group supervision does improve counselling skills in a number of different ways.
Our quality rating	Good.

Study	Ogren, M. L., & Jonsson, C. O. (2003). Psychotherapeutic skill following group supervision according to supervisees and supervisors
Item type	Journal
Journal	<i>The Clinical Supervisor</i>
Date of publication	2003
Volume	22
Issue	1
Pages	35–58
Written language	English (translated from Swedish)
Abstract	A modified Swedish version of Buckley's Self-Evaluation Scale for psychotherapeutic skills was used in this study. An extensive sample of students from three universities was investigated. Two types of scale were used – Skill and Change in skill. The supervisees rated skill both before and after supervision. Furthermore, change after supervision was rated both by supervisees and supervisors. After Principal Component analyses. Varimax rotation, three dimensions (containing emotionally loaded issues; psychodynamic understanding and mastery of working alliance) were interpreted in the supervisee ratings. However, the supervisors differentiated their ratings to a much higher degree than the supervisees. The factor structure of Skill – supervisee ratings before supervision – was used for calculating scales for all four rating formats corresponding to the three factors. There were significant increases in Skill and positive Changes of skill after group supervision. The supervisors evaluated supervisees' Changes of skill higher than the students did themselves. Significant correlations of slightly more than medium size between supervisor and supervisee ratings of Change were obtained, and this was interpreted as confirming the validity of the supervisee ratings. Skill, in the dimension Mastery of working alliance, differed between groups depending on the group supervisors according to supervisee ratings for supervision.
Type of study	Quantitative: questionnaire (pre and post), supervisee and supervisor.
Definition of supervision	None stated.
How were participants recruited?	No information was given about the methods of recruitment. The sample comprised students in postgraduate psychotherapy training courses at three Swedish Universities: Lund, Stockholm and Uppsala.
Dropouts	None stated.
Incentives	None stated.
Limitations	There are various limitations to the study. The Topics and Climate (TAC) questionnaire was developed for the study and tested on a small sample of 37 participants, which is insufficient to be sure of its reliability. One of the subscales of the TAC had poor psychometric properties. Also, Buckley's Self-Evaluation Scale (Buckley et al, 1982) was used and translated into Swedish with some items removed, which may have affected its reliability. However, both of the above studies indicate that individual and group supervision does improve counselling skills in a number of different ways. The study was also lacking in a control group; there was no comparison with a non-supervision condition.
Our quality rating	Good.

Study	Ossana, S. M. (1991). The impact of supervision on trainee self-efficacy and counseling ability: testing a developmental framework
Item type	PhD Dissertation
Publisher	U Maryland, College Park, USA.
Date of publication	1991
	<p>A test of developmental models of supervision was conducted in which the impact of congruence (ie the match of supervisory interventions with the trainee and characteristics) on supervisee self-efficacy and counselling ability was examined via regression analyses. A revision of the Supervision Level Scale (Wiley & Ray, 1986) was used to assess the developmental characteristics of supervisees and the supervision environment. Supervision dyads (n=75) from four university counselling centres participated. Data were collected at two points in time: following the third supervision session and at the conclusion of supervision from the semester. Results showed that, contrary to expectation, both congruence and incongruence were predictive of supervision outcome. The importance of congruence (and incongruence) to effective supervision was moderated by several variables including a) time (eg onset or conclusion of supervision), b) rater (eg supervisor or supervisee) and c) developmental dimension (eg beginning or advanced). Supervisors' theoretical orientation was not related to their perceptions of congruence.</p> <p>Results further revealed that trainee experience level and trainee developmental characteristics were differentially important for guiding supervisory interventions, suggesting that the developmental level and training level may not be interchangeable. In general, results suggest that the components of effective supervision are more complex than originally proposed in the developmental models of supervision. Further implications for developmental models of supervision are discussed.</p>
Type of study	Quantitative: questionnaire (pre, mid and post), supervisee and supervisor.
Definition of supervision	None provided.
How were participants recruited?	Four universities: one eastern state, one mid-western state, one western state and one private. Apart from the private one, all have comparable Masters and doctoral programs in counseling psychology. Counseling centres at three universities are involved in supervision and training of students from Masters practicum through internship. The researcher was acquainted with a staff member at each counseling centre who was willing to serve as a liaison between researcher and supervisee-supervisor dyads.
Dropouts	None stated.
Incentives	None stated.
Limitations	The study relied on supervisors' and trainees' perception to assess person-environment congruence. Since no objective measure was used, it is impossible to determine the accuracy of trainees' and supervisors' perceptions. The study relied on the use of discrepancy scores to assess person-environment congruence and changes in certain variables (eg counselling ability) across time. However, this may over simply the supervision process and neglect the presence of less prominent characteristics that may also be important.
Our quality rating	Average.

Study	Patton, M. J., & Kivlighan, D.M. (1997). Relevance of the supervisory alliance to the counseling alliance and to treatment adherence in counselor training
Item type	Journal, Article
Journal	<i>Journal of Counseling Psychology</i>
Date of publication	1997
Volume	44
Issue	1
Pages	108–115
Abstract	The quality of the supervisory working alliance should be related to the quality of the counseling working alliance and to trainee adherence to a treatment model. After each of four supervision sessions, 75 trainees completed the Supervisory Working Alliance Inventory and, after each counselling session, their clients completed the Working Alliance Inventory. Judges rated portions of the videotaped counseling sessions with an adherence measure. Hierarchical linear modelling was used to analyse these nested data (repeated measures of trainees nested within supervisors). As hypothesised, the trainee's perception of the supervisory alliance was significantly related to the client's perception of the counselling alliance and to aspects of treatment adherence. The significance of these findings is discussed in relation to the supervision literature.
Type of study	Mixed-methods: questionnaires, and observation and video taping of supervisee.
Definition of supervision	None provided.
How were participants recruited?	Clients were drawn from undergraduate pre-service, teacher education, child and adolescent development courses.
Dropouts	None stated.
Incentives	Volunteer clients received an extra credit toward their course grade as reward for participating.
Limitations	Efstation et al (1990) describe some of the limitations of this study, which include questions about statistical procedures. Also, as sessions were observed through a one-way mirror and supervisor comments were offered during the sessions, clients' perceptions of therapy may be influenced by the participation of the supervisor.
Our quality rating	Average.

Study	Raichelson, S. H., Herron, W.G., Primavera, L. H., & Ramirez, S. M. (1997). Incidence and effects of parallel process in psychotherapy supervision
Item type	Journal, Article
Journal	<i>Clinical Supervisor</i>
Date of publication	1997
Volume	15
Issue	2
Pages	37–48
Abstract	The degree to which parallel process exists in supervision and the specific effects of parallel process from the viewpoint of the participants of psychotherapy supervision were investigated. Differences between supervisors and supervisees were examined as were differences between psychoanalytic and non-psychoanalytic orientations. 300 supervisors/supervisees completed the Parallel Process Survey, which was designed specifically for this study. The results empirically validate the theoretical literature that suggests that parallel process exists, is used in psychotherapy supervision, is applicable to a variety of theoretical orientations and has clear effects on the participants of psychotherapy supervision. While there was not unanimous agreement, very few supervisors denied the existence of parallel process. Orientation differences were present, with the psychoanalytic orientation consistently reporting stronger existence and importance of parallel process and greater effects of parallel process than their non-psychoanalytic counterparts.
Type of study	Mixed methods: survey (with some qualitative analysis), supervisee and supervisor.
Definition of supervision	None stated.
How were participants recruited?	344 psychoanalytic supervisors and 148 non-psychoanalytic supervisors, plus 547 psychoanalytic supervisees and 106 non-psychoanalytic supervisees were asked to participate in the study by completing the Parallel Process Survey and a demographic data questionnaire. All supervisors worked at the post-doctoral level at institutes in New York; all supervisees were enrolled in or had completed advanced post-doctoral programs. 100 responses from the psychoanalytic supervisors and 50 responses from the non-psychoanalytic supervisors were obtained, and 100 responses from the psychoanalytic supervisees and 50 responses from the non-psychoanalytic supervisees were obtained. (No information was given as to how the responses were selected.)
Dropouts	None stated, but the full sample group did not respond to the survey.
Incentives	None stated.
Limitations	Replication studies are needed to obtain reliability and validity data on the scale properties. Further limitations – included: the representativeness of the sample, a lack of a direct measure of observed behaviour and the relatively high level of experience of the supervisee sample.
Our quality rating	Average.

Study	Steinhelber, J., Patterson, V., Cliffe, K., & Legoullon, M. (1984). An investigation of some relationships between psychotherapy supervision and patient change
Item type	Journal, Article
Journal	<i>Journal of Clinical Psychology</i>
Date of publication	1984
Volume	40
Issue	6
Pages	1346–1353
	Psychotherapy supervision has been investigated extensively, but the literature reveals few controlled studies and none that directly assess the effects of supervision on psychotherapy outcome. In this study, two aspects of psychotherapy supervision – the amount of supervision and the congruence of theoretical orientation between the supervisor and trainee – were studied in relation to patient change as measured by pre-therapy and follow-up ratings on the Global Assessment Scale (N=237). The amount of supervision was not found to be related to therapy outcome, but patients showed significantly greater improvement when their trainee-therapists reported theoretical orientations congruent with those of their supervisors.
Type of study	Quantitative: longitudinal, supervisee.
Definition of supervision	None stated.
How were participants recruited?	None stated.
Dropouts	None stated.
Incentives	None stated.
Limitations	These included the use of GAS and some concerns about its reliability; beginning and end-of-therapy GAS scores were not used; a GAS score was used at the point when the data was collected; and the therapist report was only used, (a limitation of the GAS).
Our quality rating	Average.

Study	Strozier, A. L., Kivlighan, D. M., & Thoreson, R. W. (1993). Supervisor intentions, supervisee reactions, and helpfulness – a case study of the process of supervision
Item type	Journal, Article
Journal	<i>Professional Psychology – Research and Practice</i>
Date of publication	1993
Volume	24
Issue	1
Pages	13–19
Abstract	Using an intensive single-subject design, this study examined supervisor intentions and supervisee reactions to provide information about the process of supervision. One supervisor-supervisee dyad met for 14 sessions throughout an academic semester. Each hour-long session was videotaped and reviewed using the Intentions and Reactions Lists, the Helpfulness Rating Scale and the Session Evaluation Questionnaire. Descriptive statistics and sequential analyses were used to analyse the data. The results showed that the supervisor's relationship intention was related to the supervisee's support reactions. The supervisor's Explore, Restructure, Assessment and Change intentions were related to the supervisee's Therapeutic Work reactions. Supervisee reactions were also predictive of supervisor intentions but to a lesser degree. Implications are discussed.
Type of study	Mixed methods: questionnaire and rating of video taped sessions, supervisee and supervisor.
Definition of supervision	None provided.
How were participants recruited?	No information about the recruitment of the sole supervisee. The supervisor is the second author of the study.
Dropouts	None stated.
Incentives	None stated.
Limitations	Because the authors used a case study design, the results are best seen as suggestive and heuristic. The implications that can be drawn from these results should be further investigated with group designs. The authors only examined supervisee and supervisor cognitive processes. A more complete understanding of the supervision process could be obtained by examining the interaction among various aspects of supervisor and supervisee interactions (Borders, 1991). Because the supervisee was an advanced trainee and the supervisor was experienced, results can only be meaningfully applied to trainees and supervisors at these levels. Also, because the second author was also the supervisor, his ratings (particularly helpfulness) may have been affected by his theoretical bias. There was some uncertainty as to how the sole supervisee was selected for the research. Also, the second author was supervisor and researcher and knew too much about the research methods and tools before the research began. Finally, there was too much self-reported data and no attention was paid to cultural or gender differences.
Our quality rating	Average.

Study	Tryon, G. S. (1996). Supervisee development during the practicum year
Item type	Journal, Article
Date of publication	1996
Volume	35
Issue	4
Pages	287–94
Notes	362
Abstract	This examined the validity of Stoltenberg and Delworth's (1987) Integrated Developmental Model of group supervision in a sample of 25 advanced psychotherapy practicum trainees in clinical and counseling psychology. It found longitudinal supervisee gains in self-rated self-other awareness and independence over the course of an advanced practicum year. It discusses implications for research and training.
Type of study	Quantitative: questionnaire (pre, mid and post), supervisee
How were participants recruited?	All counsellors in practicum over five years were included. No details are given of anyone who refused.
Dropouts	None stated.
Incentives	None stated.
Limitations	They give no indication of other training that students were engaged in at the time that could have influenced results. No definition of supervision is provided.
Our quality rating	Average.

Study	Vallance, K. (2004). Exploring counsellor perceptions of the impact of counselling supervision on clients
Item type	Journal, Article
Journal	<i>British Journal of Guidance and Counselling</i>
Date of publication	2004
Volume	32
Issue	4
Pages	559–574
Abstract	This qualitative study explores counsellors' experiences and perceptions of how counselling supervision impacts their clients. The literature review highlights little research in this area. Data collection combined open-ended questionnaires and semi-structured interviews. Analysis was phenomenological incorporating participant validation. The findings indicate that supervision impacts client work both helpfully and unhelpfully. Areas that emerged as having the most direct impact on client work were: exploration of client-counsellor dynamics and raising counsellor self-awareness, professional development, emotional support and the quality of the supervisory relationship. Benefits and dangers for clients not taken to supervision were highlighted. Congruence and confidence were the most direct link between supervision and client work.
Type of study	Qualitative: questionnaires and semi-structured interviews, supervisee.
Definition of supervision	None provided.
How were participants recruited?	They were known to the researcher.
Dropouts	None stated.
Incentives	None stated.
Limitations	The study has all the limitations inherent in small-scale qualitative studies. There is potential for researcher bias, as the counsellors were known to the researcher and had similar training and shared values with her. The impact on clients is only as the counsellors report it.
Our quality rating	Average.

Study	Worthen, V., & McNeill, B.W. (1996). A phenomenological investigation of “good” supervision events
Item type	Journal, Article
Journal	<i>Journal of Counseling Psychology</i>
Date of publication	1996
Volume	43
Issue	1
Pages	25–34
Notes	398
Abstract	Explored the experience of “good” supervision events from the perspective of supervisees using a qualitative phenomenological research methodology. Eight intermediate-level to advanced-level trainees participated in a tape-recorded interview in which they described a recent good supervision experience, yielding transcripts that were subjected to a qualitative analysis. A general meaning structure for the “good” supervision events was developed.
Type of study	Qualitative: semi-structured interviews with supervisees
Definition of supervision	None provided.
How were participants recruited?	Authors states willingness and availability.
Dropouts	None stated.
Incentives	None stated.
Limitations	The limitations of the study include the small sample size. As the authors note, even though the study was a phenomenological investigation to describe and uncover the structures of personal meaning, this study is not necessarily generalisable to or predictive of all good experiences of supervision and may not be inclusive of all important events or elements of what may be considered good supervision (p.33). The study only focused on the perspective of the supervisee and the authors point out that a different version of good supervision might emerge from supervisors’ perspectives. Finally, the study doesn’t link experience of good supervision with whether supervisees actually become more effective in facilitating better therapy outcomes in clients.
Our quality rating	Average.

Study	Worthington, E. L. (1984). Empirical-Investigation of Supervision of counselors as they gain experience
Item type	Journal, Article
Journal	<i>Journal of Counseling Psychology</i>
Date of publication	1984
Volume	31
Issue	1
Pages	63–75
Abstract	237 counsellors from 11 counselling agencies completed the Supervision Questionnaire, a description of the frequencies of various supervisor (155 post-PhD, 82 pre-PhD) behaviours during the preceding semester. Counsellors were classified as being in Practica 1, 2, 3, 4 or internship. Changes in supervisees' perceptions of supervisory behaviour as the counsellors gained experience were relatively congruent with previous research and theory. Supervisors do change their behaviour to match the needs of their supervisees.
Type of study	Quantitative: questionnaires with supervisee.
Definition of supervision	Yes, definition provided.
How were participants recruited?	Supervisees selected from various settings throughout the US in order to sample a variety of geographical areas.
Dropouts	None stated.
Incentives	None stated.
Limitations	Ellis and Ladany (1997) have criticised this study, saying that the problems are so severe, they obviate making inferences.
Our quality rating	Average.

6.2 Abstracts of articles/studies unable to obtain in Phase Two

Study	Alvarez, R. A. (2002). A survey and analysis of professional counselors' change in their therapeutic system through clinical supervision
Item type	PhD dissertation
Abstract	<p>The purpose of this study was to survey and analyse Licensed Professional Counselors' (LPC) change in their therapeutic system through clinical supervision. This study was concerned with determining whether the change in LPCs' therapeutic system (dependent variable) was dependent upon clinical supervision, model of clinical supervision, gender and ethnicity of the clinical supervisor, gender and ethnicity of the counselor-supervisee, the length of time the LPC was licensed, and the maturity of the LPC since attainment of licensure (independent variables). This study obtained its sample from the population of Professional Counselors licensed by the Texas State Board of Examiners of Professional Counselors from 1990 to 1999. The population sample consisted of 238 LPCs who were randomly selected. The instrument used to collect data was a Survey Questionnaire constructed by this investigator. Contents of this survey are germane to clinical supervision and were based on the literature of clinical supervision. There were eight problem statements from which eight null hypotheses were formulated and consequently tested for statistical significance at the .05 level of probability or better. All eight null hypotheses were analysed using Chi Square Test of Independence. The results of this investigation seem to show that there is a significant dependence (association) between LPCs' change in therapeutic system and the following independent variables: clinical supervision, model of clinical supervision, gender and ethnicity of the clinical supervisor, and gender and ethnicity of the professional counsellor. Statistical significance was found at the .01 level of probability. There appears to be no dependence (association) between change in counsellors' therapeutic system and length of time the LPC was licensed. Also, there seems to be no dependence between the change in counsellors' therapeutic system and the maturation of the LPC since attainment of licensure. There was no statistical significance found at the .05 level of probability. The results of this study provided research-based data from clinical supervision, not clinical training, thus avoiding definition confusion, which is a barrier to attainment of good standard in clinical supervision practice. The results also provided research-based data regarding the role(s) of gender and ethnicity in clinical supervision.</p>

Study	Christenberry, N. J., & Jones, C. H. (1990). Impact of faculty supervisors on counseling trainees' attitudes toward social influence
Item type	Report
Date of publication	1990
Abstract	<p>Counsellors should be able to increase their effectiveness by appropriate use of social influence strategies; learning to use power bases is a critical need of counsellor trainees. One way they can learn to use power bases is by modelling their supervisors in counseling practica. In this study, the effects of faculty supervisors' attitudes on those of counselling trainees towards power bases were examined. Supervisors' and trainees' attitudes were measured using a questionnaire developed by Robyak (1981) for assessing referent, legitimate and expert power base preferences. During a one-semester counselling practicum, pre/post measures of preference strength were obtained from 50 trainees in eight supervisory groups. Analysis of variance procedures yielded no significant differences between repeated measures or between groups. Observed trends consistent with desirable use of power bases indicated that trainees were not unduly influenced by their supervisors' preferences. Implications of these results for counselor training programs are discussed and references are included. (Author/TE) Notes 9 p., Paper presented at the Annual Meeting of the Mid-South Educational Research Association (New Orleans, LA, November 14, 1990).</p>

Study	Gorman, J. (1989). Counsellor supervision: accounting for change or no-change from a systems perspective
Item type	Report
Publisher	Publisher: La Trobe University, Bundoora Vic, 1989.
Date of publication	1989
Abstract	This paper presents a systems analysis of the effects that supervision has on post-supervisory counselling sessions. Using material gathered from audio taped counselling and supervision sessions, the interventions that counsellors use before and after supervision are compared. The research questions which the paper addresses are firstly whether there is any change in counsellor interventions following supervision and, secondly, how one would account for either change or no-change in counsellor interventions from a systems perspective. Ideas and explanations are drawn from the literature, with particular reference to the works of Bateson and his cybernetic epistemology; White and his application to therapy of Bateson's theories; Watzlawick and others of the Mental Research Institute who have applied Group Theory and the Theory of Logical Types to therapy; Ricci, Selvini-Palazzoli and their colleagues from the Milan school who discuss 'therapeutic neutrality', and Keeney who considers cybernetic epistemology from an aesthetic perspective.

Study	Grothaus, T. J. P. (2004). An exploration of the relationship between school counselors' moral development, multicultural counseling competency and their participation in clinical supervision
Item type	PhD dissertation
Date of publication	2004
Abstract	The purpose of this study was to investigate the relationship between school counselors' moral development level, their self-perceived multicultural counseling competence level and their level of participation in clinical supervision. The professional literature suggests that school counsellors are facing enhanced job complexity and demands in increasingly diverse school communities. The literature also suggests that school counsellors' skills and performance are more likely to decline than improve over the years. In addition, there is a dearth of research concerning the professional development of school counsellors. Given the support in the literature regarding the benefits of participation in clinical supervision, the theoretical frameworks of moral development and multicultural counseling were utilised to examine the provision of clinical supervision as a means of assisting school counsellors in meeting these complex challenges and opportunities. The findings did not support the hypotheses for this study. The relationship between participation in clinical supervision and enhanced moral development and increased levels of multicultural counselling competence was not statistically significant. Also, no statistically significant relationship was discovered between the instruments employed to measure moral development (Defining Issues Test, Short Version) and self-perceived multicultural counselling competency (Multicultural Counseling Knowledge and Awareness Scale). Implications of the findings and suggestions for further research were discussed.

Study	Nguyen, T. V. (2004). A comparison of individual supervision and triadic supervision
Item type	PhD dissertation
Date of publication	2004
Abstract	<p>This study was designed to measure and compare individual supervision to triadic supervision in promoting counsellor effectiveness and development. During individual supervision, one counsellor met with one supervisor for an hour. Two models of triadic supervision were created for this study: Split Focus and Single Focus. Triadic consists of two supervisees and one supervisor meeting for one hour. During the Split Focus, 30 minutes was allocated to each counsellor for supervision. During the Single Focus, the whole hour was spent supervising only one of the counsellors. The next week, the whole hour was spent supervising the other counsellor. Three comparison groups were employed to determine the effectiveness of the three supervision models. An instrument was used to evaluate counsellor effectiveness and another instrument was used to evaluate counsellor development. Forty-seven Masters-level counselling students enrolled in practicum participated in this study. The practicum met for 16 weeks. Each counselor filled out a Supervisee Levels Questionnaire-Revised at the beginning (pre-test) and at the end (post-test) of the semester. This instrument determined the counsellor's developmental growth. Each counselor submitted a tape of a counselling session at the beginning (pre-tape) and at the end (post-tape) of the semester. The tape was rated on-site by the doctoral supervisor utilising the Counselor Rating Form-Short. An objective rater also appraised the submitted tapes utilising the same instrument. The instrument determines counsellor effectiveness. At the end of the study, an Analysis of Covariance determined that the three supervision models did differ in developmental growth. The Split Focus grew significantly compared to Single Focus and compared to individual supervision. However, the Single Focus grew significantly on the factor of self and other awareness compared to individual. In terms of effectiveness, an Analysis of Covariance determined that the three supervision models did not differ significantly.</p>

Item	Romei, J. (2004). Internalization and clinical development in psychotherapy supervision: study of psychologists-in-training
Item type	PhD dissertation
Abstract	<p>The aim of this study was to investigate trainee internalisation of his or her supervisor and the development of clinical skills over time in the context of the supervisory relationship. Internalisation of the supervisor by the trainee may be one mechanism by which clinical growth is achieved in supervision, akin to the manner in which patients change in psychotherapy (Geller & Farber, 1993). Literature has posited that neophyte trainees are more likely to internalise their supervisors than master therapists (Hanlon, 1990), while clinical growth is thought to occur in a variety of areas including self and other awareness, motivation to be a therapist, and fluctuating dependency and autonomy vis-a-vis their supervisors (Stoltenberg, 1997). Such internalisation and clinical development is thought to occur within the context of a strong supervisory working alliance (Efstation, Patton & Kardash, 1990). The sample consisted of 34 pre-doctoral level trainees and 16 supervisors. Both trainees and supervisors completed questionnaires assessing for demographic data. Trainees completed the Supervisor Representation Inventory, Supervisee Levels Questionnaire-Revised, Supervision Working Alliance Inventory-Trainee Version, Supervision Experience Questionnaire, and the Psychotherapy Metaphors Questionnaire at the start and end of a six-month time period, equivalent to the supervision over an academic year. Supervisors completed the Psychotherapy Metaphors Questionnaire and the Supervision Working Alliance Inventory-Supervisor Version at the start of the academic year. Trainee internalisation and clinical development were thought to increase over time, while the supervisory working alliance acted as a covariate. Due to the low number of supervisors participating in this study, supervisor measures were excluded from analyses. Findings indicated a decrease in both the formal properties of trainee internalisations as well as a significant decrease in trainee involvement with thoughts regarding supervision, contrary to expectations. Formal properties were thought to be associated with affective components of internalisation and, as such, related more to the affective relationship of psychotherapy than the functional relationship of supervision. The decrease in trainee involvement with thoughts about their supervisors was thought to be related to anxious rumination or worry concerning supervision and competence as a therapist.</p>

Study	Stoltenberg, C., & Holmes, D. (1988). A comparison of counseling supervisors' and trainees' perceptions of development
Item type	Report
Date of publication	1988
Abstract	Stoltenberg's (1981) counselor complexity model, a developmental model of the supervision process, conceptualises the training process as a sequence of four identifiable stages through which a trainee progresses: dependency, dependency-autonomy conflict, conditional dependency and the master counsellor. Evidence has been found in support of developmental models of supervision in general and of Stoltenberg's constructs specifically. This study investigated the degree of difference or agreement between supervisors' perceptions of trainees and their trainees' perceptions of themselves. Twenty pairs of supervisors and trainees rated the trainees on their perceived developmental level by completing the Supervisee Levels Questionnaire during the third week of the practicum and again at the end of the semester. A correlational analysis indicated a positive relationship between the trainees' ratings of satisfaction with supervision and their perception of developmental growth over the course of the semester. The results indicated that the larger the difference between the pre-test and the post-test ratings, the greater the degree of satisfaction reported by the trainees. In addition, the greater the difference between the trainees' and the supervisors' initial perception of the trainees' development, the greater the subsequent satisfaction with the supervisor. (NB) Notes 15 p., Paper presented at the Annual Meeting of the Southwestern Psychological Association (34th, Tulsa, OK, April 21–23, 1988).

Study	Trembath, S. A. (1988). A serendipitous process? Immediate effects of supervisor interventions on trainee behaviour
Item type	Report
Publisher	Publisher: La Trobe University, Bundoora Vic, 1988.
Date of publication	1988
Abstract	Not many people seem to know in detail what happens in supervision sessions, if literature is any guide. Actual trainee supervisor interactions are rarely investigated. By means of a content analysis of counselling sessions before and after supervision, and of the supervision session itself, the present research illustrates how, if at all, trainees follow up suggestions or leads given by supervisors. An attempt is also made of a new synthesis of the existing research regarding supervisor styles, roles and behaviours. Lists of supervisor behaviours are categorised according to the three major supervisor roles of counsellor, teacher and consultant. These lists form the framework for analysing supervisor interventions in the research component of the present study. The question being investigated is: what immediate effects do supervisor interventions have on the counselling behaviour of trainees? Previous studies into the effectiveness of supervision are critically reviewed, highlighting this author's rationale for a small-scale, neutral investigation. Process and outcome dimensions are analysed using six supervisor-trainee-client triads as subjects. A series of audiotapes and transcripts were made for each triad over three sessions and analysed according to the themes and goals of supervision, supervisor roles and focus, evidence of discrepant reporting and outcome measures of those aspects of trainees' behaviour pertinent to the goals of supervision. From the content analysis, the immediate effects of supervisors' teaching, counselling and consulting are far from obvious. The author makes conclusions and recommendations to supervisors aimed at reducing the apparently serendipitous process of supervision, at least as revealed in the cases in this study.

6.3 Overviews of relevant supervision research reviews

Ellis, M. V., Ladany, N., Kregel, M., & Schult, D. (1996). Clinical supervision research from 1981 to 1993: a methodological critique. *Journal of Counseling Psychology*, 43(1): 35–50.

Abstract

Investigated empirical studies in clinical supervision published from 1981 through to 1993 to assess scientific rigor and test whether the quality of methodology had improved since the review by Russell, Crimmings and Lent (1984). Data revealed a shift to realistic field studies, unchecked Type I and Type II error rates, medium Effect Sizes and inattention to hypothesis validity. (KW)

General aim/review question

To assess the status of clinical supervision research via a thorough documentation of the major methodological validity threats and to determine the extent to which supervision researchers have replicated previous studies.

Inclusion criteria

Pre-practicum training was excluded.

Included:

- Met definition of clinical supervision
- Published 1981–93
- Supervision of individual counselling
- Group supervision of individual counselling
- Case studies
- Peer supervision
- It included case studies, school, rehabilitation, community, mental health counselling, counselling and clinical psychology, social work, psychiatry and psychiatric nursing.
- It included peer supervision.
- It excluded pre-practicum supervision, teacher supervision, speech therapy and supervision of group therapy.

How many studies? 130 articles and one book – 144 studies.

How is it different to our review? The focus of the review was methodology and methodological limitations and errors.

Are some of their studies included in our review? Yes, but only with respect to methodological problems.

What was the conclusion of the review? That more methodological rigour is needed.

Ellis, M. V., & Ladany, N. (1997). Inferences concerning supervisees and clients in clinical supervision. In C. E. Watkins Jr. (Ed.), *Handbook of Psychotherapy Supervision*. New York, Wiley: 447–507.

Abstract

No abstract

General aim/review question

To address the limitations of the review by Ellis et al. (1996) by conducting a more focused review of supervision research that incorporated a) a systematic evaluation of the scientific rigour of each study, b) reinterpreted findings and 3) organised review according to inferences.

Inclusion criteria

Inferences are defined as rudimentary core premises and tentative conclusions about supervision, supervisees or clients. That the study:

- 1 addressed supervision of individual counselling or therapy;
 - 2 empirically based article in journal from 1981;
 - 3 focused on inferences about supervisee or client variables.
- It included case studies, school, rehabilitation, community, mental health counselling, counselling and clinical psychology, social work, psychiatry and psychiatric nursing.
 - It included peer supervision.
 - It excluded pre-practicum supervision, teacher supervision, speech therapy and supervision of group therapy.

How many studies? 95 research articles and one book, a total of 104 studies.

How is it different to our review?

Inclusion criteria was much broader. Two of the six areas of inference overlapped with our review: inferences regarding supervisee development and client outcome. The focus of the review was an intense and thorough methodological critique.

Are some of their studies included in our review?

Yes, the studies in the above sections (approximately 10 studies).

What was the conclusion of the review?

That the majority of studies were methodologically flawed. That a lot can be learned from the mistakes of others. That there are some promising instruments that might inform future research. That it highlighted the gaps in what is reliably known about supervision.

Freitas, G. J. (2002). The impact of psychotherapy supervision on client outcome: a critical examination of 2 decades of research. *Psychotherapy*, 39: 4, 354–367.

Abstract

The supervision of counselors-in-training has become an important area of psychotherapy research. Through supervision, one expects that trainees will gain the knowledge and clinical skills needed to work effectively with clients. Much research has been dedicated to understanding trainee skill development through supervision, but relatively few studies in the literature have addressed whether clinical supervision leads to improved client outcome. The current article presents a review of two decades of research into the question of whether clinical supervision improves client outcome. Ten studies conducted between 1981 and 1997 are reviewed in detail. Caveats and criticisms, as well as suggested directions for future research, are discussed.

General aim: review question

To address in detail the 10 studies between 1981 and 1993 that reported on client outcome and supervision.

Inclusion criteria

Nine studies identified by Ellis et al. (1996) plus three others, reduced to 10 in total. Three were excluded because two were case studies limited in validity and one was primarily about training.

How many studies?

10.

How is it different to our review?

Inclusion criteria broader. Studies not all focused on clinical counselling. Methodological critique of each study in detail, which focuses exclusively on client outcome. Latest study 1996.

Are some of their studies included in our review?

Yes. The methodological weaknesses are noted.

What was the conclusion of the review?

Methodological problems lead to limited conclusions.

Wheeler, S. (2003). *Research on supervision of counsellors and psychotherapists: a systematic scoping search*. Rugby, BACP.

Abstract

The aim of this review is to provide a comprehensive report on supervision research published in the English language between 1980 and 2002. No research articles have purposely been excluded from the report because of size of sample, methodology and context of supervision if there is some analysis of the findings. Descriptive, philosophical, anecdotal and case study articles have been excluded.

General aim/review question

The aim of this review is to provide a comprehensive report on supervision research published in the English language between 1980 and 2002.

Inclusion criteria

Supervision related to a broad range of professions was included.

The following were excluded: books and discursive articles that did not report empirical research and management supervision.

No of studies: 388.

How is it different to our review?

It is much broader than the current review with much looser exclusion criteria. The review provides references organised in categories:

Models of supervision:

- B The experience of supervision
- C Supervisory relationship
- D Events in supervision
- E Supervision process
- H Ethical considerations
- J Supervision of trainees
- K Supervision of experienced practitioners
- L Training of supervisors
- M Supervision mode
- N Cross-cultural issues in supervision
- O Effectiveness of supervision
- R Roles, tasks and function
- T Supervision in other professions
- V Gender issues in supervision
- Y Characteristics of the supervisor

No attempt was made to analyse the studies in depth and only brief details of methodology and sample size are given.

Are some of their studies included in our review? Yes, studies that meet the inclusion criteria for the current review are included.

What was the conclusion of the review? The review concluded that there were many opportunities for future research into supervision that takes account of the methodological difficulties encountered in previous work. It provides a convenient reference document to quickly identify the scope of existing studies that has the potential to enable future researchers.

Whittaker, S. M. (2004). *A multi-vocal synthesis of supervisees' anxiety and self-efficacy during clinical supervision meta-analysis and interviews*. PhD dissertation, Virginia Polytechnic Institute and State University.

Abstract

Clinical supervision of counsellors in training is an integral part of their professional and personal development. Accrediting bodies in academia and licensure standards in most fields require beginning professionals to receive clinical supervision. During clinical supervision, supervisees frequently experience anxiety, and the supervisees' self-efficacy, or belief about their ability to counsel clients, is affected by supervision. The questions addressed in this study were to what extent does clinical supervision affect supervisees' anxiety and self-efficacy and do different types of supervision have varying effects on their anxiety and self-efficacy?

A meta-analysis comprised of 10 studies was conducted to determine the influence of supervision on supervisees' anxiety and self-efficacy. Clinical supervision was found to have a medium effect, $ES=.454$ and $ES=.430$, on supervisees' anxiety. Clinical supervision had a large effect, $ES=.655$, on supervisees' self-efficacy. In addition, a qualitative review of the studies included in the meta-analysis yielded methodological concerns in the areas of adequate control group, sample size, representativeness of sample and follow-up assessment.

Due to the small number of studies meeting the meta-analysis criteria, quantitative findings were limited. Therefore, individual interviews with clinical supervisors and supervisees were conducted to corroborate or refute the findings of the meta-analysis and to lend multiple "voices" in an attempt to answer the research questions. Face-to-face interviews were conducted with nine supervisees and five supervisors in a Counselors' Education program. The results of the interviews corroborated the finding of the meta-analysis that clinical supervision affects supervisees' anxiety and self-efficacy by decreasing both. All types of supervision were described as increasing anxiety and self-efficacy with no particular type predominating.

Limitations of the research and implications for educators, practitioners and future research are discussed. A limitation of the meta-analysis was the relatively small number of existing studies meeting the criteria for inclusion. This limited the interpretation of the findings in terms of answering the research questions. The interview portion of the research was limited due to the use of a purposive sample, participants all being students from the same program and the researcher was also a student in this program.

General aim: review question

To examine the extent to which clinical supervision affects supervisees' anxiety and self-efficacy and to assess whether different types of supervision, ie individual supervision, group supervision and live supervision, have varying effects on supervisees' anxiety and self-efficacy.

Inclusion criteria

Each study must have investigated the effect of some type of clinical supervision on supervisees' anxiety and/or self-efficacy. Types of clinical supervision included individual supervision, live supervision and group supervision as the author defines. Triadic supervision was not included. All types and techniques of clinical supervision were included.

How many studies?

Meta-analysis on 10 studies and interviews with supervisors and supervisees.

How is it different to our review?

Of the 10 studies included in the meta-analysis, four were published in refereed journals and six were doctoral dissertations. The four journal articles have been reviewed for the systematic review, but it has not been possible to obtain the dissertations. The four articles are:

Daniels, J. A., & Larson, L. M. (2001). The impact of performance feedback on counseling self-efficacy and counsellor anxiety. *Counselor Education and Supervision*, 41: 120–131.

Ellis, M. V., Krenzel, M., & Beck, M. (2002). Testing self-focused attention theory in clinical supervision. Effects on supervisee anxiety and performance. *Journal of Counseling Psychology*, 49: 101–106.

Johnson, M. E. (1989). Effects of self-observation and self-as-a model on counsellor trainees' anxiety and self-evaluations. *The Clinical Supervisor*, 7: 59–70.

Williams, E., Judge, A. B., Hill, C. E., & Hoffman, M. A. (1997). Experiences of novice therapists in pre-practicum: trainees', clients' and supervisors' perceptions of therapists' personal reactions and management strategies. *Journal of Counseling Psychology*, 44: 390–399.

None of these four articles have been included in this systematic review because they do not meet the inclusion criteria. In each, case studies had an experimental design that used role-play clients, simulated client material or volunteer clients who were recruited for the purpose of the study for one short session only as part of counsellor training.

Are some of their studies included in our review?

No

What was the conclusion of the review?

Clinical supervision had:

- a medium to large effect on supervisee anxiety and self-efficacy
- a medium effect on anxiety during supervision and a large effect on counselling self-efficacy during supervision.

Individual supervision had:

- a medium effect on supervisee anxiety
- a low to medium effect on supervisee self-efficacy.

Peer group supervision had:

- a medium effect on supervisee anxiety (one study).

Live supervision had:

- a low effect on anxiety (one study)
- a low effect on size of group supervision of self-efficacy (one study).

No conclusion can be reached in this meta-analysis as to the different types of supervision's effect on self-efficacy other than individual supervision.

Section 7: References

7.1 Articles included in the review

- Bambling, M., King, R., Raue, P., Schweitzewr, & Lambert, W. (2006). Clinical supervision: Its influence on client-rated working alliance and client rated working alliance and client symptom reduction in the brief treatment of major depression. *Psychotherapy Research*, 16 (3): 317–331.
- Borders, L. D. (1990). Developmental changes during supervisees' first practicum. *Clinical Supervisor*, 8 (2): 157–167.
- Carey, J. C., Williams, K. S., & Wells, M. (1988). Relationships between Dimensions of Supervisors' Influence and Counselor Trainees' Performance. *Counselor Education and Supervision*, 28: 130–139.
- Cashwell, T. H., & Dooley, K. (2001). The impact of supervision on counselor self-efficacy. *Clinical Supervisor Special Issue*: 20 (1): 39–47.
- Couchon, W. D., & Bernard, J. M. (1984). Effects of timing of supervision on supervisor and counselor performance. *The Clinical Supervisor*, 2(3): 3–20.
- Dodenhoff, J. T. (1981). Interpersonal attraction and direct-indirect supervisor influence as predictors of counsellor trainee effectiveness. *Journal of Counseling Psychology*, 28 (1):47–52.
- Earley, T. M. (2004). Counselor supervision post-licensure: Relationship of clinical supervision type and frequency to supervision anxiety and counselor self-efficacy. PhD Dissertation: Capella University, USA.
- Efstation, J. F., Patton, M. J., & Kardash, M. (1990). Measuring the working alliance in counsellor supervision. *Journal of Counseling Psychology*, 37 (3): 322–329.
- Guest, P. D., & Beutler, L. E. (1988). Impact of Psychotherapy Supervision on Therapist Orientation and Values. *Journal of Consulting & Clinical Psychology*, 56(5): 653–58.
- Jiménez Andújar, D. J. (2002). *Experiences of personal growth of master's-level supervisees within the supervisory process*. PhD Dissertation: University of New Orleans.
- Jumper, S. A. (1999). *Immediate feedback using the 'bug-in-the-ear' in counselor training: Implications for counseling self-efficacy, trainee anxiety and skill development*. PhD Dissertation: University of North Dakota.
- Ladany, N., Ellis, M. V., & Friedlander, M. L. (1999). The supervisory working alliance, trainee self-efficacy and satisfaction. *Journal of Counseling and Development*, 77: 447–455.
- Leherman-Waterman, D., & Ladany, N. (2001). Development and validation of the evaluation process within supervision inventory. *Journal of Counseling Psychology*. 48 (2): 168–177.
- Milne, D. L., Pilkington, J., Gracie, J., & James, I. (2003). Transferring skills from supervision to therapy: A qualitative and quantitative N=1 analysis. *Behavioural and Cognitive Psychotherapy*, 31: 193–202.
- Ogren, M. L., & Jonsson, C.O. (2003). Psychotherapeutic skill following group supervision according to supervisees and supervisors. *The Clinical Supervisor*, 22(1): 35–58.
- Ogren, M. L., Jonsson, C. O., & Sundin, E. C. (2005). Group supervision in psychotherapy. The relationship between focus, group climate and perceived attained skill. *Journal of Clinical Psychology*, 61 (4): 373–388.

Ossana, S. M. (1990). *The impact of supervision on trainee self-efficacy and counseling ability: Testing a developmental framework*. PhD Dissertation: University of Maryland College, USA.

Patton, M. J., & Kivlighan, D. M. (1997). Relevance of the supervisory alliance to the counseling alliance and to treatment adherence in counselor training. *Journal of Counseling Psychology*, 44 (1): 108–115.

Raichelson, S. H., Herron, W. G., Primavera, L. H., & Ramirez, S. M. (1997). Incidence and effects of parallel process in psychotherapy supervision. *The Clinical Supervisor*, 15(2): 37–48.

Steinheber, J., Patterson, V., Cliffe, K., & LeGoullon, M. (1984). An investigation of some relationships between psychotherapy supervision and patient change. *Journal of Clinical Psychology*, 40 (3): 1346–1353.

Strozier, A. L., Kivlighan, D. M., & Thoreson, R. W. (1993). Supervisor intentions, supervisee reactions and helpfulness: A case study of the process of supervision. *Professional Psychology: Research and Practice*, 24 (1): 13–19.

Tryon, G. S. (1996). Supervisee Development during the Practicum Year. *Counselor Education & Supervision* 1996, 35 (4): 287–94.

Vallance, K. (2004). Exploring counsellor perceptions of the impact of counselling supervision on clients. *British Journal of Guidance and Counselling*. 32(4): 559–574.

Worthen V., & McNeill, B. W. (1996). A phenomenological investigation of 'good' supervision events. *Journal of Counseling Psychology*, 43 (1): 25–34.

Worthington, E. L. (1984). Empirical investigation of supervision of counselors as they gain experience. *Journal of Counseling Psychology*, 31 (1): 63–75.

7.2 Related supervision research reviews

Ellis, M. V., & Ladany, N. (1997). Inferences concerning supervisees and clients in clinical supervision. In C. E. Watkins Jr (Ed.), *Handbook of Psychotherapy Supervision* (p.447–507). New York: Wiley.

Ellis, M. V., Ladany, N., Kregel, M., & Schult, D. (1996). Clinical Supervision Research from 1981 to 1993: A Methodological Critique. *Journal of Counseling Psychology* 43(1): 35–50.

Freitas, G. J. (2002). The impact of psychotherapy supervision on client outcome: A critical examination of 2 decades of research. *Psychotherapy*, 39: 4, 354–367.

Milne, D., & James, I. (2000). A systematic review of effective cognitive-behavioural supervision. *British Journal of Clinical Psychology* 39 (Pt 2): 111–127.

Wheeler, S. (2003). Research on supervision of counsellors and psychotherapists: a systematic scoping search. Rugby, BACP.

Whittaker, S. M. (2004). A multi-vocal synthesis of supervisees' anxiety and self-efficacy during clinical supervision meta-analysis and interviews. PhD Dissertation: Virginian Polytechnic and State University.

7.3 Articles included in Phase Two but unobtainable

Alvarez, R. A. (2002). A survey and analysis of professional counselors' change in their therapeutic system through clinical supervision. Dissertation Abstracts International Section A: Humanities and Social Sciences. (Full reference unknown)

- Christenberry, N. J., & Jones, C. H. (1990). Impact of Faculty Supervisors on Counseling Trainees' Attitudes toward Social Influence Strategies. Report. (Full reference unknown)
- Gorman, J. (1989). Counsellor supervision: accounting for change or no-change from a systems perspective. Report: Australia: La Trobe University, Victoria. (Full reference unknown)
- Grothaus, T. J. P. (2004). An exploration of the relationship between school counselors' moral development, multicultural counseling competency, and their participation in clinical supervision. Dissertations Abstracts International: Section A: Humanities and Social Sciences. (Full reference unknown)
- Nguyen, T. V. (2004). A comparison of individual supervision and triadic supervision. Dissertations Abstracts International: Section A: Humanities and Social Sciences. (Full reference unknown)
- Romei, J. (2004). Internalization and clinical development in psychotherapy supervision: A study of psychologists-in-training. Dissertations Abstracts International: Section B: The Sciences and Engineering. (Full reference unknown)
- Stoltenberg, C., & Holmes, D. (1988). A Comparison of Counseling Supervisors' and Trainees' Perceptions of Development. Paper presented at the Annual Meeting of the Southern Western Psychological Association (34th Tulsa, OK, April 21–23).
- Trembath, S. A. (1988). A serendipitous process? Immediate effects of supervisor interventions on trainee behaviour. Report: Australia: La Trobe University, Victoria. (Full reference unknown)
- 7.4 Articles excluded during Phase Two**
- Arundale, C. J. (1993). Psychotherapy supervision: impacts, practices and expectations. PhD, University College London, UK.
- Ashby, R. H. (1999). Counselor development and supervision: An exploratory study of the Intergrated Developmental Model of supervision. PhD Dissertation: University of Oklahoma, USA.
- Baker, B. G. (2003). The process of group supervision: a grounded theory of masters students' experiences. PhD Dissertation: Idaho State University, USA.
- Birk, J. M., & Mahalik, J.R. (1996). The influence of trainee conceptual level, trainee anxiety and supervision evaluation on counselor developmental level. *Clinical Supervisor*, 14 (1): 123–137.
- Blocher, D., Christensen, E., Hale-Fiske, R., Neren, S., Spencer, T., & Fowlkes, S. (1985). Development and preliminary validation of an instrument to measure cognitive growth. *Counselor Education and Supervision*, 25 (1): 21–30.
- Bowen, N. H. (1990). The effects of supervisee training level and supervisor education level on the state anxiety, trait anxiety, and self-rating skill of counselors-in-training. PhD Dissertation: University of Nebraska, Lincoln, USA.
- Calley, N. G. (2001). Sequential supervisory styles: examining the impact on counselor development. PhD Dissertation: Wayne State University.
- Constantine, M. G. (2001). Multiculturally-focused counseling supervision: Its relationship to trainees' multicultural counseling self-efficacy. *The Clinical Supervisor*, 20 (2): 81–90.
- Cox, K. C. (2000). The in-session inner experiences of a supervisor with her two trainees: A qualitative study. Unpublished PhD Dissertation: Loyola, University of Chicago.
- Dougher, M. K. (2001). Multiple concurrent supervision: The impact on case conceptualization. Dissertations Abstracts International: Section B: The Sciences & Engineering. PhD Dissertation: university unknown.
- Hanscom, K. L. (1990). The relationship between counselor conceptual level and supervisory style in the supervision of brief psychotherapy. PhD Dissertation: University of Maryland College Park, USA.
- Humeidan, M. A. (2002). Counseling self-efficacy, supervisory working alliance, and social influence in supervision. PhD Dissertation: university unknown.
- Kaupp, M. A. (2001). A phenomenological exploration of the trainee's experience with computer-assisted live supervision. PhD Dissertation: university unknown.
- Kilminster, S. M., & Jolly, B. C. (2000). Effective supervision in clinical practice settings: a literature review. *Medical Education*, 34 (10): 827–840.
- Koob, J. J. (2002). The Effects of Solution-Focused Supervision on the Perceived Self-Efficacy of Therapists in Training. *Clinical Supervisor*, 21 (2): 161–183.
- Krasner, R. F., Howard, K. I., & Brown, A. S. (1998). The acquisition of psychotherapeutic skill: an empirical study. *Journal of Clinical Psychology*, 54 (7): 859–903.
- Ladanay, N., Constantine, M. G., & Hofheniz, E. W. (1997). Supervisee multicultural case conceptualisation ability and self-reported multicultural competence as functions of supervisees racial identity and supervisor focus. *Journal of Counseling Psychology*, 44 (284–293): 411.
- Ladany, N., Brittan-Powell, C. S., & Pannu, R. K. (1997). The Influence of Supervisory Racial Identity Interaction and Racial Matching on the Supervisory Working Alliance and Supervisee Multicultural Competence. *Counselor Education and Supervision*, 36 (4): 284–304.
- Ladany, N., Inman, A. G., Constantine, M. G., & Hofheinz, E. W. (1997). Supervisee multicultural case conceptualization ability and self-reported multicultural competence as functions of supervisee racial identity and supervisor focus. *Journal of Counseling Psychology*, 44 (3): 284–293.
- Lambert, M. (1997). Research and the Supervisory Process. In Hess, A. K. (Ed.), *Psychotherapy supervision: theory, research and practice* (pp. 423–451). USA: J Wiley & Sons.
- Lambert, M. J., & Hawkins, E. J. (2001). Using information about patient progress in supervision: Are outcomes enhanced? *Australian Psychologist Special Issue: Training in clinical and counseling psychology*, 36 (2): 131–138.
- Lambert, M. J., & Ogles, B. M. (1997). The effectiveness of psychotherapy supervision. In Watkins, E. C. (Ed.), *Handbook of psychotherapy supervision* (pp.421–446). USA: John Wiley & Sons.
- Larson, L. M., & Daniels, J. A. (1998). Review of the counseling self-efficacy literature. *Counseling Psychologist*, 26 (2): 179–218.
- Lenihan, G. O., & Kirk, W. G. (1992). Conjoint supervision with beginning trainees: The model and its effectiveness. *Clinical Supervisor*, 10 (1): 35–50.
- Malikisilozos, M., Mehnert, W. O., Work, G. G., & Gold, J. (1981). Differential Supervision and Cognitive Structure Effects on Empathy and Counseling Effectiveness. *International Journal for the Advancement of Counselling*, 4 (2): 119–129.
- Martin, G. E., & McBride, M. C. (1987). The Results of the Implementation of a Professional Supervision Model on Counselor Trainee Behavior. *Counselor Education and Supervision*, 27 (2): 155–67.

Milne, D. (1999). Evidence-based clinical supervision. *Clinical Psychology Forum*, 133 (November): 32–36.

Milne (in press). An empirical definition of clinical supervision. *British Journal of Clinical Psychology*.

Rowland, N. (2007). BACP and NICE. A Submission to a House of Commons Select Committee inquiry into NICE (the National Institute for Health and Clinical Excellence). *Therapy Today*, Vol. 18 (5): 27–30.

Sandell, R. (1985). Influence of supervision, therapist's competence and patient's ego level on the effects of time-limited psychotherapy. *Psychotherapy and Psychosomatics*, 44: 103–109.

Sells, J. N., Goodyear, R. K., Lichtenberg, J. W., & Polkinghorne, D. E. (1997). Relationship of Supervisor and Trainee Gender to In-Session Verbal Behavior and Ratings of Trainee Skills. *Journal of Counseling Psychology*, 44 (4): 406–412.

Shanklin, J. E. (2003). *Supervision and increasing self-efficacy in the therapist-trainee* PhD Dissertation: Biola University, USA.

Silvestri, T. J. (2003). The temporal effect of supervisor focus, the supervisory working alliance, and the graduate training environment upon supervisee multicultural competence. PhD Dissertation: university unknown.

Singo, W. E. (1998). The effects of peer group supervision and individual supervision on the anxiety, self-efficacy, and basic skill competency of counselor trainees in practicum. PhD Dissertation: Wayne State University, USA.

Sommer, C. A. (2004). Vicarious traumatization and the role of supervision: An exploration of sexual violence counselors' stories. PhD Dissertation: university unknown.

Strauss, D. M. (1995). The relationship between supervisory factors and counseling self-efficacy factors. PhD Dissertation: University of Missouri, Saint Louis, USA.

Sumerel, M. B., & Borders, L.D. (1996). Addressing Personal Issues in Supervision: Impact of Counselors' Experience Level on Various Aspects of the Supervisory Relationship. *Clinical Supervisor*, 13 (1): 91–100.

Tagger, L. S. (2003). The relationship of collective efficacy and performance in supervision. PhD Dissertation: University unknown.

Thurber, S. L. (2001). Analysis of the supervisor-therapist-client system via computer assisted live supervision of marital therapy. PhD Dissertation: Brigham Young University.

Torres-Rivera, E. Phan, L. T., Maddux, C., Wilbur, M. P., & Garrett, M. T. (2001). Process versus content: Integrating personal awareness and counseling skills to meet the multicultural challenge of the twenty-first century. *Counselor Education and Supervision*, 41 (1): 28–40.

Tracey, T. J., & Sherry, P. (1993). Complementary interaction over time in successful and less successful supervision. *Professional Psychology: Research and Practice*, 24(3): 304–311.

Ward, J. E. (2001). An exploration of the development of counseling self-efficacy during the supervision process: including the dynamics of counselor development, the supervisory relationship, satisfaction with supervision, and experience. PhD Dissertation: Southern Illinois University.

Weber, S. A. (2001) Evaluating the ability of students to articulate case conceptualizations: A qualitative study of two methods of group supervision. PhD Dissertation: University unknown.

West, W. C. (2004). Learnings from a qualitative study into counselling supervision: listening to supervisor and supervisee. *Counselling and Psychotherapy Research*, 4 (2): 20–26.

Whittaker, S. M. (2004) A multi-vocal synthesis of supervisees' anxiety and self-efficacy during clinical supervision meta-analysis and interviews. PhD Dissertation: Virginian Polytechnic and State University.

Wiggins, J. D., & Moody, A. (1983). Identifying Effective Counselors through Client-Supervisor Ratings and Personality-Environmental Variables. *Vocational Guidance Quarterly*, 13 (2): 259–269.

Winter, M., & Holloway, E. L. (1991). Relation of trainee experience, conceptual level, and supervisor approach to selection of audiotaped counseling passages. *Clinical Supervisor*, 9 (2), 87–103.

Worthington, E. L. (1987). Changes in Supervision as Counselors and Supervisors Gain Experience – a Review. *Professional psychology – Research and Practice*, 18 (3): 189–208.

Young, J. S., & Borders, L. D. (1999). The intentional use of metaphor in counseling supervision. *Clinical Supervisor*, 18 (1): 137–149.

4.5 Articles from Phase One Screening

If details of any of the other excluded references are required, these are available directly from the authors. Many of the over 8,000 abstracts screened did not bear relevance to either research or the supervision of counselling and psychotherapy.

Section 8: All references

- BACP. (2003). *The Accreditation of Counsellor Training Courses*. Rugby: BACP.
- BACP. (2002). *Ethical framework for good practice in counselling and psychotherapy*. Rugby: BACP.
- Bahrck, A. S. (1990). Role induction for counsellor trainees: Effects the supervisory working alliance. *Dissertation Abstracts International*, 51, 1484B (University Microfilms No 90-14, 392).
- Bambling, M., King, R., Raue, P., Schweitzewr, & Lambert, W. (2006). Clinical supervision: Its influence on client-rated working alliance and client rated working alliance and client symptom reduction in the brief treatment of major depression. *Psychotherapy Research*, 16 (3): 317-331.
- Barak, A., & LaCrosse, M. B. (1975). Multidimensional perception of counsellor behaviour. *Journal of Counseling Psychology*, 22: 471-476.
- Beck, A., Steer, R., & Garbin, M. (1987). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, 8, 77-100.
- Bernard, J. M., & Goodyear, R. K. (1992). *Fundamentals of Clinical Supervision*. London: Pearson.
- Bernard, J. M., & Goodyear, R. K. (2004). *Fundamentals of Clinical Supervision*. London: Pearson.
- Borders, L. D. (1990). Developmental changes during supervisees' first practicum. *Clinical Supervisor*, 8 (2): 157-167.
- Buckley, P., Conte, H., Plutchik, R., Karasu, T., & Wild, K. (1982). Learning dynamic psychotherapy: A longitudinal study. *American Journal of Psychiatry*, 139: 1607-1611.
- Cochrane Collaboration (2006). <http://www.cochrane.org/resources/handbook/Handbook4.2.6Sep2006.pdf>.
- Carey, J. C., Williams, K. S., & Wells, M. (1988). Relationships between Dimensions of Supervisors' Influence and Counselor Trainees' Performance. *Counselor Education and Supervision*, 28: 130-139.
- Carroll, M. (1996). *Counselling Supervision: Theory, Skills and Practice*. London, Cassell.
- Cashwell, T. H., & K. Dooley (2001). The impact of supervision on counselor self-efficacy. *Clinical Supervisor Special Issue*: 20(1): 39-47.
- Couchon, W. D., & Bernard, J. M. (1984). Effects of timing of supervision on supervisor and counselor performance. *The Clinical Supervisor*, 2 (3): 3-20.
- Deeks, J. J., Dinnes, J., D'Amico, R., Sowden, A. J., Skarovitch, C., & Song, F. (2003). Evaluating non-randomised intervention studies. *Health Technology Assessment*, 7 (27).
- Dodenhoff, J. T. (1981). Interpersonal attraction and direct-indirect supervisor influence as predictors of counsellor trainee effectiveness. *Journal of Counseling Psychology*, 28 (1): 47-52.
- Earley, T. M. (2004). Counselor supervision post-licensure: Relationship of clinical supervision type and frequency to supervision anxiety and counselor self-efficacy. PhD Dissertation: Capella University, USA.
- Efstation, J. F., Patton, M. J., & Kardash, M. (1990). Measuring the working alliance in counsellor supervision. *Journal of Counseling Psychology*, 37 (3): 322-329.
- Ekstein, R., & Wallerstein, R (1972). *The teaching and learning of psychotherapy*. New York: Int. University Press.
- Elliott, R. (1985). Helpful and non-helpful events in brief counselling interviews: An empirical taxonomy. *Journal of Counseling Psychology*, 32: 307-322.
- Ellis, M. V., Ladany, N., Kregel, M., & Schutt, D. (1996). Clinical Supervision Research from 1981 to 1993: A Methodological Critique. *Journal of Counseling Psychology*, 43 (1): 35-50.
- Ellis, M. V., & Ladany, N. (1997). Inferences concerning supervisees and clients in clinical supervision. In C. E. Watkins (Ed.), *Handbook of Psychotherapy Supervision* (pp. 447-507). New York: Wiley.
- Endicott, J., Spitzer, R. L., Fleiss, J. L., & Cohen, J. (1976). The Global Assessment Scale. *Archives of General Psychiatry*, 33: 766-771.
- EPPI (2006a). <http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=67>
- EPPI (2006b). <http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=68>
- Eysenck, H. J., & Eysenck, S. B. G. (1969). *Personality structure and measurement*. San Diego, CA: Knapp.
- Friedlander, M. L., & Synder, J. (1983). *Self-efficacy inventory: Correlation with experience level*. Unpublished raw data.
- Friedlander, M. L., & Ward, L. G. (1984). Development and validation of the supervisory styles inventory. *Journal of Counseling Psychology*, 31: 541-557.
- Freitas, G. J. (2002). The impact of psychotherapy supervision on client outcome: a critical examination of two decades of research. *Psychotherapy: Theory, Research, Practice, Training*, 39: 354-367.
- Guest, P. D., & L. E. Beutler (1988). Impact of Psychotherapy Supervision on Therapist Orientation and Values. *Journal of Consulting & Clinical Psychology* 56 (5): 653-58.
- Hansen, J. C., Robins, T. H., & Grimes, J. (1982). Review of research on practicum supervision. *Counselor Education and Supervision*, 22 (1): 15-23.
- Hawkins, P. & Shohet, R. (1989). *Supervision in the Helping Professions*. Buckingham: Open University Press.
- Hawkins, P., & Shohet, R. (2000). *Supervision in the Helping Professions*. 2nd edition Buckingham: Open University Press.
- Henderson, C. E., Cawyer, Stringer, C., & Watkins, C. E., Jr. (1999). A comparison of student and supervisor perceptions of effective practicum supervision. *Clinical Supervisor*, 18 (1): 47-74.
- Heppner, P. O., & Handley, P. (1981). A study of interpersonal influence process in supervision. *Journal of Counseling Psychology*, 28: 437-444.
- Heppner, P. O., & Handley, P. (1982). The relationship between supervisory behaviours and perceived supervisor expertness, attractiveness, or trustworthiness. *Counselor Education and Supervision*, 22, 37-41.
- Hess, A. K. (1987). Psychotherapy supervision: Stages, Buber, and a theory of relationship. *Professional Psychology: Research and Practice*, 18: 251-259.
- Higgins, J. P. T., & Green, S., (Eds.) (2005). Cochrane Handbook for Systematic Reviews of Interventions 4.2.5 [updated May 2005]. <http://www.cochrane.org/resources/handbook/hbook.htm> (accessed 28 September 2005).
- Hill, C. E., Helms, J. E., Spiegel, S. B., & Tichenor, V. (1988). Development of a system for categorizing client reactions to therapist interventions. *Journal of Counseling Psychology*, 35: 27-36.

- Hill, C. E., & O'Grady, K. E. (1985). List of therapist intentions illustrated in a case study and with therapist of varying theoretical orientations. *Journal of Counseling Psychology*, 32: 3–22.
- Holloway, E. L. (1995). *Clinical Supervision – A Systems Approach*. Thousand Oaks: Sage.
- Holloway, E. L., & Wampold, B. E. (1984). Dimensions of satisfaction in the supervision interview. *Paper presented at the meeting of the American Psychological Association Convention*: Toronto, Canada.
- Horvarth, A. O., & Greenberg, L. (1989). Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology*, 36: 223–232.
- Inskipp, F., & Proctor, B. (1995). *Becoming a Supervisor* (1st Edition). Twickenham: CASCADE.
- Inskipp, F., & Proctor, B. (2001a). *Making the most of Supervision* (2nd Edition). London: Cascade.
- Inskipp, F., & Proctor, B. (2001b). *Becoming a Supervisor* (2nd Edition). Twickenham: CASCADE.
- Jacobs, D., David, P., & Meyer, D. J. (1995). *The Supervisory Encounter*. New Haven: Yale University Press.
- Jiménez Andújar, D. J. (2002). *Experiences of personal growth of master's-level supervisees within the supervisory process*. PhD Dissertation: University of New Orleans.
- Johnson, E., Baker, S. B., Kopala, M., Kiselica, M. S., & Thompson, E. C. (1989). Counseling self-efficacy and counseling competence in prepracticum training. *Counselor Education and Supervision*, 28: 205–218.
- Jumper, S. A. (1999). Immediate feedback using the 'bug-in-the-ear' in counselor training: Implications for counseling self-efficacy, trainee anxiety and skill development. PhD Dissertation: University of North Dakota.
- Kadushin, A. (1968). Games people play in supervision. *Social Work*, 13 (July): 23–32.
- Kiesler, D. J. (1993). Supervision in interpersonal communication psychotherapy. In A. K. Hess (Ed.), *Psychotherapy supervision: Theory, research and practice* (Vol.2). New York: Wiley.
- Kilminster, S. M., & B. C. Jolly (2000). Effective supervision in clinical practice settings: a literature review. *Medical Education*, 34 (10): 827–840.
- Ladany, N., Ellis, M. V., & Friedlander, M. L. (1999). The supervisory working alliance, trainee self-efficacy and satisfaction. *Journal of Counseling and Development*. 77: 447–455.
- Lambert, M. J., & Ogles, B. M. (1997). The effectiveness of psychotherapy supervision. In C. Watkins (Ed.), *Handbook of psychotherapy supervision* (pp. 421–446). New York, NY, USA, John Wiley & Sons Inc.
- Larson, L. M., & Daniels, J. A. (1998). Review of the counseling self-efficacy literature. *The Counseling Psychologist*, 26 (2): 179–219.
- Larson, L. M., Suzuki, L. A., Gillespie, K. N., Potenza, M. T., Bechtel, M. A., & Toulouse, A. L. (1992). Development and validation of the counselling self-estimate inventory. *Journal of Counseling Psychology*, 39: 105–120.
- Leach, M. M., Stoltenberg, C. D., McNeill, B. W., & Eichenfield, G. A. (1997). Self-efficacy and counselor development: Testing the integrated developmental model. *Counselor Education & Supervision*, 37 (2): 115–124.
- Leherman-Waterman, D., & Ladany, N. (2001). Development and validation of the evaluation process within supervision inventory. *Journal of Counseling Psychology*. 48 (2): 168–177.
- Loganbill, C., Hardy, E., & Delworth, U. (1982). Supervision: A conceptual model. *The Counseling Psychologist*, 10: 3–42.
- Loganbill, C., & Hardy, E. (1983). Developing training programmes for clinical supervisors. *The Clinical Supervisor*, 1:15–21.
- Mattinson, J. (1977). *The Reflection Process in Casework Supervision*. London: Tavistock Institute of Human Relations.
- McClure, W. J. (1973). The absentee-cuing of counselor in training: An evaluation of extended and abbreviated cues and various effects of the system upon clients, counsellors and trainers. Unpublished doctoral dissertation, Wayne State University.
- McNeill, B. W., Stoltenberg, C. D., & Pierce, R. A. (1985). Supervisees' perceptions of their development: A test of the counsellor complexity model. *Journal of Counseling Psychology*, 32: 630–633.
- McNeill, B. W., Stoltenberg, C. D., & Romans, J. S. C. (1992). The Integrated Developmental Model of Supervision – Scale Development and Validation Procedures. *Professional Psychology-Research and Practice*, 23(6): 504–508.
- Melchert, T. P., Hays, V. L., Wiljanen, L. M., & Koleck, A. K. (1996) Testing models of counselor development with a measure of Counselor Self-efficacy. *Journal of Counseling and Development*, 74: 640–644.
- Milne, D. (in press). An empirical definition of clinical supervision. *British Journal of Clinical Psychology*.
- Milne, D., & James, I. (2000). A systematic review of effective cognitive-behavioural supervision. *British Journal of Clinical Psychology*. 39 (Pt 2): 111–27.
- Milne, D. L., Pilkington, J., Gracie, J., & James, I. (2003). Transferring skills from supervision to therapy: A qualitative and quantitative N=1 analysis. *Behavioural and Cognitive Psychotherapy*, 31:193–202.
- Milton Keynes Primary Care Trust. (2002). Critical appraisal skills programme (CASP): Making sense of evidence. Milton Keynes PCT.
- Myrick, R. D., & Kelly, F. D., Jr. (1971). A scale for evaluating practicum students in counseling and supervision. *Counselor Education and Supervision*, 10: 330–336.
- Oakley, A., Gough, D., Oliver, S., & Thomas, J. (2005). The politics of evidence and methodology: lessons from the EPPI-Centre. *Evidence & Policy: A Journal of Research, Debate and Practice*, 1 (1): 5–31.
- Ogren, M. L., & Jonsson, C. O. (2003). Psychotherapeutic skill following group supervision according to supervisees and supervisors. *The Clinical Supervisor*, 22 (1): 35–58.
- Ogren, M. L., Jonsson, C. O., & Sundin, E. C. (2005). Group supervision in psychotherapy. The relationship between focus, group climate and perceived attained skill. *Journal of Clinical Psychology*, 61 (4): 373–388.
- Olsson, G. (1996). Självkänsla i obalans: En studie över terapeutelevens uppfattning om sina fardigheter (Self-esteem out of balance: A study of therapy students' comprehension of their skills. *Nordisk psykologi*, 48 (4): 252–265.
- Orlinsky, D., & Rønnestad, H. (2005). *How Psychotherapists Develop*. New York: American Psychological Association.

- Ossana, S. M. (1990). The impact of supervision on trainee self-efficacy and counseling ability: Testing a developmental framework. PhD Dissertation: University of Maryland College, USA.
- Page, S., & Wosket, V. (2001). *Supervising the counsellor: A cyclical model* (2nd Edition). PA, Brunner-Routledge: Philadelphia.
- Patton, M. J., & Kivlighan, D. M. (1997). Relevance of the supervisory alliance to the counseling alliance and to treatment adherence in counselor training. *Journal of Counseling Psychology*, 44 (1): 108–115.
- Raichelson, S. H., Herron, W. G., Primavera, L. H., & Ramirez, S. M. (1997). Incidence and effects of parallel process in psychotherapy supervision. *The Clinical Supervisor*, 15(2): 37–48.
- Robiner, W. N., & Schofield, W. (1990). References on supervision in clinical and counseling psychology. *Professional Psychology – Research & Practice*, 21(4): 297–312.
- Rokeach, M. (1973). *The nature of human values*. New York: Free Press.
- Rokeach, M. (1979). *Understanding human values*. New York: Free Press.
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied*, 80 (1, Whole No.609).
- Rowland, N. (2007). BACP and NICE: A submission to a House of Commons Select Committee Inquiry into NICE (National Institute for Health and Clinical Excellence). *Therapy Today*, 18, (5): 27–30.
- Russell, R. K., Crimmings, A. M., & Lent, R. W. (1984). Counselor training and supervision. In S. Brown & R. Lent (Eds.), *Handbook of Counseling Psychology*. New York: Wiley.
- Searles, H. F. (1955). *The informational value of the supervisors' experience. Collected Papers on Schizophrenia and Related Subjects*. London, Hogarth Press.
- Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). Manual for the State-Trait Anxiety Inventory. *Psychological Bulletin*, 124: 240–261.
- Steinheilber, J., Patterson, V., Cliffe, K., & LeGoullon, M. (1984). An investigation of some relationships between psychotherapy supervision and patient change. *Journal of Clinical Psychology*, 40 (3): 1346–1353.
- Stiles, W. B., & Snow, J. S. (1984). Counseling session impact as viewed by novice counsellors and their clients. *Journal of Counseling Psychology*, 31: 3–12.
- Stoltenberg, C. (1981). Approaching supervision from a developmental perspective: the counselor complexity model. *Journal of Counseling Psychologists*, 28: 59–65.
- Stoltenberg, C. D., & Delworth, U. (1987). *Supervising counselors and therapists*. San Francisco: Jossey Bass.
- Storrow, H. A. (1960). The measurement of outcome of psychotherapy. *Archives of General Psychiatry*, 1: 142–146.
- Strozier, A. L., Kivlighan, D. M., & Thoreson, R. W. (1993). Supervisor intentions, supervisee reactions and helpfulness: A case study of the process of supervision. *Professional Psychology: Research and Practice*. 24 (1): 13–19.
- Sundland, D. M. (1997a June). Theoretical orientation. A multi-professional, American Sample. Paper presented at the meeting of the Society for Psychotherapy Research, Madison, WI.
- Sundland, D. M. (1977b). Theoretical orientations of psychotherapists. In Gurman A. S. & Razon A. M. (Eds.), *Effective psychotherapy: A handbook of research* (pp. 189–222). New York: Pergamon Press.
- Tracey, T. J., & Sherry, P. (1993). Complementary interaction over time in successful and less successful supervision. *Professional Psychology: Research and Practice*, 24(3): 304–311.
- Tryon, G. S. (1996). Supervisee Development during the Practicum Year. *Counselor Education & Supervision* 1996, 35 (4): 287–94.
- Vallance, K. (2004). Exploring counsellor perceptions of the impact of counselling supervision on clients. *British Journal of Guidance and Counselling*. 32(4): 559–574.
- Wheeler, S. (2003). *Research on supervision of counsellors and psychotherapists: a systematic scoping search*. Rugby: BACP.
- Whittaker, S. M. (2004) A multi-vocal synthesis of supervisees' anxiety and self-efficacy during clinical supervision meta-analysis and interviews. PhD Dissertation: Virginian Polytechnic and State University.
- Wiley, M. O., & Ray, P. B. (1986). Counseling supervision by developmental level. *Journal of Counseling Psychology*, 33: 439–445.
- Worthen, V. E., & Lambert, M. J. (2007). Outcome oriented supervision: advantages of adding systematic client tracking supportive consultations. *Counselling and Psychotherapy Research*, 7(1): 48–53.
- Worthen, V., & McNeill, B. W. (1996). A phenomenological investigation of 'good' supervision events. *Journal of Counseling Psychology*, 43 (1): 25–34.
- Worthington, E. L. (1987). Changes in Supervision as Counselors and Supervisors Gain Experience – a review. *Professional Psychology-Research and Practice*, 18(3): 189–208.
- Worthington, E. L., Jr., & Roehlke, H. J. (1979). Effective supervision as perceived by beginning counselors-in-training. *Journal of Counseling Psychology*, 26, 64–73.

Appendix: Data Extraction Form

Section A: Administration

A.1 Name of reviewer	A.1.1 Sue Wheeler A.1.2 Kaye Richards A.1.3 Sophia A.1.4 Other (1) A.1.5 Other (2) A.1.6 Other (3)
A.2 Date of review	A.2.1 Date <i>Specify</i>
A.3 Abstract/summary of key findings	A.3.1 Not included A.3.2 Included
A.4 Keywords <i>(as stated in article)</i>	A.4.1 Keywords not included A.4.2 Keywords included (specify)
A.5 Country of in which study took place <i>N.B. This is not necessarily the same as the country of the research institutions. If the study is in more than one country, indicate this.</i>	A.5.1 USA A.5.2 Australia A.5.3 Canada A.5.4 UK A.5.5 Other (please specify)
A.6 Source of funding	A.6.1 Not stated A.6.2 Stated (specify)

Section B: Aims and Context of the Study

B.1 What were the aims of the study? <i>Please write in author's description if there is one. Elaborate if necessary but indicate which is reviewer's interpretation.</i>	B.1.1 Not stated/unclear (specify) B.1.2 Explicitly stated (specify) B.1.3 Implicit (specify)
B.2 Does this research follow on from a previous study?	B.2.1 Not stated/unclear (specify) B.2.2 Explicitly stated (specify) B.2.3 Implicit (specify)
B.3 Research setting of the counselling	B.3.1 Private practice B.3.2 Government department B.3.3 FE institution B.3.4 HE institution (specify if undergraduate or postgraduate) B.3.5 Secondary school B.3.6 Primary school B.3.7 Rehabilitation centre B.3.8 Psychiatry B.3.9 Nursing B.3.10 Hospital – other than nursing /psychiatry (specify) B.3.11 Company/workplace B.3.12 Clinic (specify) B.3.13 Other (specify) B.3.14 Doesn't state
B.4 Type of supervision	B.4.1 Work with real clients B.4.2 Skills training/development B.4.3 Both clients and training B.4.4 Other (specify)
B.5 Set-up of research	B.5.1 Naturalistic – supervision already existed B.5.2 Supervision set up for the purpose of study B.5.3 Other (specify)
B.6 Context of supervision	B.6.1 Individual B.6.2 Pairs B.6.3 Group B.6.4 Online B.6.5 Telephone B.6.6 Other (specify) B.6.7 Not stated

B.7 Type of counselling being supervised	B.7.1 Individual B.7.2 Couple B.7.3 Individual – telephone B.7.4 Individual – online B.7.5 Other (specify) B.7.6 Not stated
B.8 Theoretical model of therapy	B.8.1 Humanistic B.8.2 Cognitive and behavioural B.8.3 Psychodynamic B.8.4 Integrative B.8.5 Eclectic B.8.6 Other (specify) B.8.7 Not stated
B.9 Duration of supervision (specify)	B.9.1 Specifically organised for the project (give details) B.9.2 Ongoing

Section C: Biographical Information of Supervisor(s)

C.1 Years of experience as a therapist	C.1.1 No experience C.1.2 Less than 5 years C.1.3 5-10 years C.1.4 11-15 years C.1.5 More than 15 years C.1.6 Unknown
C.2 Years of experience as a supervisor	C.2.1 Less than 5 years C.2.2 5-10 years C.2.3 10-15 years C.2.4 More than 15 years C.2.5 Unknown
C.3 Ethnicity	C.3.1 Stated (give details) C.3.2 Not stated
C.4 Gender (give details)	C.4.1 Male (give numbers) C.4.2 Female (give numbers) C.4.3 Not stated
C.5 Age (give details)	C.5.1 Details (specify) C.5.2 Not stated
C.6 Total number of supervisor(s) in study (specify)	C.6.1 Details
C.7 Theoretical model of therapy	C.7.1 Psychodynamic C.7.2 Integrative C.7.3 Cognitive behavioural C.7.4 Humanistic C.7.5 Eclectic C.7.6 Stated (specify) C.7.7 Not stated
C.8 Theoretical model of supervision	C.8.1 Not stated C.8.2 Not explicitly stated <i>Write in as worded by reviewer</i> C.8.3 Stated <i>Write in as written by author</i>
C.9 Supervisor training (specify)	C.9.1 Not stated C.9.2 Stated (give details)
C.10 Aspect of supervisor studied (specify)	C.10.1 Style C.10.2 Interventions C.10.3 Behaviour C.10.4 Ethnicity C.10.5 Personality C.10.6 Other (specify)
C.11 How is the supervision dimension measured/studied?	C.11.1 Give details C.11.2 N/A

C.12 How clear is the description of the model of supervision used?	C.12.1 Not at all C.12.2 A little C.12.3 Satisfactory C.12.4 Good C.12.5 Excellent
C.13 Other relevant information (give details)	C.13.1 Details

Section D: Biographical Information of Supervisee(s)

D.1 Age	D.1.1 Stated (specify) D.1.2 Not stated
D.2 Gender	D.2.1 Male D.2.2 Female D.2.3 Mixed gender (give details of breakdown) D.2.4 Not stated
D.3 Ethnicity	D.3.1 Not stated D.3.2 Stated (give details)
D.4 Training (theoretical orientation)	D.4.1 Psychodynamic D.4.2 Humanistic D.4.3 Cognitive behavioural D.4.4 Integrative D.4.5 Eclectic D.4.6 Other (specify) D.4.7 Not stated
D.5 Years of counselling experience	D.5.1 0–2 yrs D.5.2 2–4 yrs D.5.3 5–10 yrs D.5.4 More than 10 yrs D.5.5 Other (specify) D.5.6 Not stated
D.6 Theoretical orientation of therapy that supervision relates to	D.6.1 Psychodynamic D.6.2 Humanistic D.6.3 Cognitive behavioural D.6.4 Integrative D.6.5 Eclectic D.6.6 Other (specify) D.6.7 Not stated
D.7 Is the supervisee a trainee?	D.7.1 Yes D.7.2 No D.7.3 Not stated
D.8 Has the supervisee(s) been trained in using supervision?	D.8.1 Yes (specify) D.8.2 No D.8.3 Not stated
D.9 Other relevant details	D.9.1 Details

Section E: Clients in the study

E.1 How many clients in counselling are involved in the study?	E.1.1 Number (specify) E.1.2 Not stated
E.2 What is the gender of clients?	E.2.1 Male (numbers) E.2.2 Female (numbers) E.2.3 Not stated
E.3 Presenting problem of client(s)	E.3.1 Known (give details) E.3.2 Not known
E.4 Consent to research requested?	E.4.1 Yes E.4.2 No E.4.3 Not stated
E.5 What is the ethnicity of clients?	E.5.1 Known (give details) E.5.2 Not known

E.6 Any other details of clients?	E.6.1 Yes (give details) E.6.2 No
-----------------------------------	--------------------------------------

Section F: Research Methods and Design

F.1 Does the study address an appropriate and clearly focused question?	F.1.1 Yes (identify question) F.1.2 No (identify question) F.1.3 Maybe (identify question)
F.2 What interventions or factors are evaluated in the study? <i>(please specify)</i>	F.2.1 Give details
F.3 Is a definition of supervision given?	F.3.1 Yes (state how referenced/to who it is attributed) F.3.2 No
F.4 Sampling and recruitment procedures	F.4.1 Details
F.5 Were any incentives provided to recruit people into the study?	F.5.1 Yes (specify) F.5.2 No F.5.3 Not stated
F.6 Were participants asked for consent before entering the study?	F.6.1 No, not stated F.6.2 Unclear F.6.3 Requested from participant F.6.4 Requested from others (specify)
F.7 What proportion of the sample consented for the study?	F.7.1 Not stated F.7.2 Unclear F.7.3 Stated (specify)
F.8 Was the study piloted? <i>This does not include similar interventions tested by others</i>	F.8.1 Yes, previously piloted with a sample of target population (specify) F.8.2 Not stated F.8.3 Unclear F.8.4 Authors consider this to be pilot F.8.5 Yes, previously piloted with others (specify)
F.9 Do the authors indicate any specific barriers to developing/delivering the intervention?	F.9.1 Yes (specify) F.9.2 No
F.10 Do the authors indicate any factors favourable to developing/delivering the intervention?	F.10.1 Yes (specify) F.10.2 Not stated
F.11 Was any information provided on those who dropped out of the study?	F.11.1 Unclear F.11.2 Not relevant F.11.3 Yes, reported (specify) F.11.4 No
F.12 Who collected the research data?	F.12.1 Supervisor F.12.2 Counsellor F.12.3 Researcher (author) F.12.4 Researcher (not the author) F.12.5 Other (specify) F.12.6 Not stated F.12.7 Unclear
F.13 Was special training provided for any of the supervisors in relation to the project?	F.13.1 Not stated F.13.2 Unclear F.13.3 Yes (specify) F.13.4 No
F.14 What impact(s) on the supervisee were examined? <i>List all the impacts studied</i>	F.14.1 Self-efficacy F.14.2 Self-confidence F.14.3 Competence F.14.4 Case conceptualisation F.14.5 Multicultural competence F.14.6 Skills F.14.7 Ethical practice F.14.8 Decision making F.14.9 Assessment skill F.14.10 Motivation F.14.11 Professional identity F.14.12 Job satisfaction F.14.13 Knowledge F.14.14 Outcome for patient F.14.15 Others (specify) F.14.16 No impact measured

F.15 Any other relevant aspects of the study?	F.15.1 No F.15.2 Yes (specify)
F.16 What general approach was used to collect data?	F.16.1 Not stated F.16.2 Unclear F.16.3 Only quantitative (specify) F.16.4 Only qualitative (specify) F.16.5 Mixed-methods (specify)
F.17 What methods were used to collect data? <i>For all the methods used please provide any relevant information about tools/approach used as part of data collection</i>	F.17.1 Not stated F.17.2 Unclear F.17.3 Documentation F.17.4 Focus groups – supervisee F.17.5 Focus groups – supervisor F.17.6 Focus groups – client F.17.7 Interviews – supervisee F.17.8 Interviews – supervisor F.17.9 Interviews – client F.17.10 Observation – of supervisee F.17.11 Observation of – supervisor F.17.12 Observation – of client F.17.13 Self-completion report/diary – supervisee F.17.14 Self-completion report/diary – supervisor F.17.15 Self-completion report/diary – client F.17.16 Standardised q'aire – supervisee (specify q'aire) F.17.17 Standardised q'aire – supervisor (specify q'aire) F.17.18 Standardised q'aire – client (specify q'aire) F.17.19 Other q'aire – supervisee (specify q'aire) F.17.20 Other q'aire – supervisor (specify) F.17.21 Other q'aire – client (specify) F.17.22 Video recording – supervisee F.17.23 Video recording – supervisor F.17.24 Video recording – client F.17.25 Online data F.17.26 Others (specify)
F.18 When did the collection of data take place?	F.18.1 Not stated F.18.2 Unclear F.18.3 Pre and post <i>Give any dates</i> F.18.4 Post only <i>Give any dates</i> F.18.5 Concurrently F.18.6 For a limited period during intervention <i>Give any dates</i> F.18.7 Other (specify)
F.19 How are the participants allocated to intervention and control/ comparison groups?	F.19.1 Random – information given (specify) F.19.2 Intervention and control comparison group F.19.3 Non random (specify) F.19.4 Random (no information given) F.19.5 Not stated F.19.6 Not relevant (study not a trial)
F.20 Was the allocation to intervention and control comparison groups done blind?	F.20.1 Not relevant (study not a trial) F.20.2 Not stated F.20.3 Unclear (specify) F.20.4 Yes F.20.5 No
F.21 Were participants aware which group they were in for this research?	F.21.1 Not relevant (study not a trial) F.21.2 Not stated F.21.3 Unclear (specify) F.21.4 Yes F.21.5 No
F.22 Was outcome measurement done blind? <i>ie were those assessing the outcomes aware whether the participants had been in a control comparison group or intervention group?</i>	F.22.1 Not relevant (study not a trial) F.22.2 Not stated F.22.3 Unclear (specify) F.22.4 Yes F.22.5 No

<p>F.23 What was the attrition or participation rate? <i>On the basis of those from whom baseline data were collected (make it clear whether it is attrition or participation that is reported).</i></p>	<p>F.23.1 Reported for study population as a whole (specify) F.23.2 Reported for one/some group(s) (specify) F.23.3 Reported for all groups (or for study population as a whole if only one group) (specify) F.23.4 Unclear F.23.5 Not stated F.23.6 Not relevant</p>
---	--

Section G: Results, Discussions and Conclusions

<p>G.1 Are there any other useful results? (give a BRIEF overview) <i>Only give a brief overview of any other results so we know what other issues are evident in studies.</i></p>	<p>G.1.1 Yes (specify) G.1.2 No</p>
<p>G.2 What are the results that relate to the impact of supervision on the supervisee?</p>	<p>G.2.1 Give details</p>
<p>G.3 Do you consider there are any limitations of the study that the author(s) have not identified?</p>	<p>G.3.1 No G.3.2 Yes (specify)</p>
<p>G.4 Do the authors state any limitations of the study?</p>	<p>G.4.1 Details</p>
<p>G.5 Are there any obvious inconsistencies in the reporting of data?</p>	<p>G.5.1 Yes (specify) G.5.2 No G.5.3 Not sure</p>
<p>G.6 Does the data presented substantiate the author's findings?</p>	<p>G.6.1 No (specify) G.6.2 Unclear G.6.3 Yes</p>
<p>G.7 Is there agreement between the authors and the reviewer?</p>	<p>G.7.1 Yes (specify) G.7.2 No (specify)</p>
<p>G.8 Is this study important to the future development of the use of supervision? <i>State whether it is important to research and/or practice developments</i></p>	<p>G.8.1 No G.8.2 Yes, definitely (in what way?) G.8.3 Yes, somewhat (in what way?) G.8.4 Not sure (specify) G.8.5 Probably not</p>
<p>G.9 Is the study replicable from the report? <i>A clear description of research design and process.</i></p>	<p>G.9.1 The research design is replicable G.9.2 The research design is not replicable G.9.3 The supervision intervention is replicable G.9.4 The supervision intervention is not replicable</p>
<p>G.10 How methodologically sound is this study?</p>	<p>G.10.1 1 – Not G.10.2 2 – Dubious (why?) G.10.3 3 – Satisfactory G.10.4 4 – Good G.10.5 5 – Excellent</p>
<p>G.11 How well does this study fit the inclusion criteria?</p>	<p>G.11.1 1 – Not at all G.11.2 2 – Dubious (say more) G.11.3 3 – A little (say more) G.11.4 4 – Somewhat G.11.5 5 – A lot</p>
<p>G.12 Rate the overall quality of the research.</p>	<p>G.12.1 1 – Very poor G.12.2 2 – Poor G.12.3 3 – Average G.12.4 4 – Very good G.12.5 5 – Excellent</p>