

# How Does Clinical Supervision Work? Using a “Best Evidence Synthesis” Approach to Construct a Basic Model of Supervision

Derek Milne  
Helen Aylott  
Helen Fitzpatrick  
Michael V. Ellis

**ABSTRACT.** Clinical supervision is of growing importance, but poor conceptualization continues to impede research and practice. We conducted a “best evidence synthesis” (empirical review) of the literature to generate an integrative summary of the concepts and models used (implicitly or explicitly) in 24 published empirical articles. Using the qualitative data from these adequately rigorous, successful manipulations of supervision, we constructed inductively a basic model of clinical supervision. That is, we proffer an evidence-based, conceptually

---

Derek Milne is affiliated with Newcastle University, Newcastle-upon-Tyne, England, and Northumberland, Tyne and Wear NHS Trust.

Helen Aylott is affiliated with the School of Psychology, Newcastle University, Newcastle-upon-Tyne, England.

Helen Fitzpatrick is affiliated with the School of Psychology, Newcastle University, Newcastle-upon-Tyne, England.

Michael V. Ellis is affiliated with the School of Education, State University of New York at Albany, Albany, NY.

Address correspondence to Derek Milne, Doctorate in Clinical Psychology, 4th Floor, Ridley Building, Newcastle University, Newcastle-upon-Tyne NE1 7RU, England. E-mail: d.l.milne@ncl.ac.uk

The Clinical Supervisor, Vol. 27(2) 2008

© 2008 by The Haworth Press. All rights reserved.

doi: 10.1080/07325220802487915

integrative, and suitably complex model of supervision. The model features 32 contextual variables of successful supervision (e.g., administrative support), 26 supervision interventions (i.e., corrective feedback), and 28 outcomes (i.e., how supervisees learned from supervision).

**KEYWORDS.** Clinical supervision, conceptual model, systematic review

One of the major challenges hindering the further development of clinical supervision has been poor conceptualization in both theory and empirical studies: Clarifying what clinical supervision is and how it works. Supervision theories have historically been derived from applying models of therapy, from extrapolating models from other fields, and from clinical experience (Leddick & Bernard, 1980). Few if any current models or theories of clinical supervision were derived at least in part from the empirical supervision literature. Although dozens of supervision models have been proposed (Bernard & Goodyear, 2004), most lack conceptual rigor, fail to recognize its complexity (Watkins, 1995a; Worthington, 1987), rarely focus on the actual behaviors of supervisors and supervisees in the context of moderating environmental factors (Watkins, 1997), and have at best moderate empirical validity (Ellis & Ladany, 1997). Hence, conceptualizing current supervision remains problematic.

To illustrate, Ellis, Ladany, Krengel, and Schult (1996) reviewed the clinical supervision literature and found that at least 80% of the 144 studies reviewed were poorly conceptualized (e.g., implicit theorizing, inconsequential or ambiguous hypotheses, and conceptual and methodological flaws). Although 20% of the studies explicitly tested conceptualizations, the majority were internally inconsistent (i.e., there was a mismatch between the theorizing and the methodology). As a result, the capacity to test supervision theory was "severely compromised in nearly every study" (Ellis et al., 1996, p. 44). This was seen as part of a gradual shift to atheoretical pragmatic field studies. Like many other authors (e.g., Chen, 1990; Kerlinger & Lee, 2000; Serlin, 1987; Tracey & Glidden Tracey, 1999), Ellis et al. argued that studies should be

“explicitly theory-driven” (p. 45). Specifically, researchers need to explicate and test theory and delineate the presumed relations among constructs, providing a lucid rationale that logically relates the constructs to one another and explains how they interrelate, so that unambiguous hypotheses can be set out and tested. For example, this could take the form of a path diagram that shows how the model operates (see Harkness, 1997).

There is, therefore, an understandable disquiet with current models of supervision. Ellis and Ladany (1997) noted that neither the conceptual rigor nor the empirical evidence lend support. They suggested that clinical supervision is a more complex phenomenon than depicted in current models. Consequently, they argued that they need to be suitably revised and made interpretable within empirical research. Falender and Shafranske (2004) agreed: “an empirical, evidence-based theoretical foundation is required” (p. 232). In short, the problem is multifold: The majority of clinical supervision models and theories are untested and not grounded in the empirical literature, most empirical investigations do not explicate and test supervision theory, and much of the extant empirical literature lacks conceptual-methodological rigor (e.g., Bernard & Goodyear, 2004; Ellis & Ladany, 1997; Holloway & Neufeldt, 1995; Stein & Lambert, 1995).

### ***HOW CAN WE MAKE PROGRESS?***

A promising strategy to construct an empirically based model of clinical supervision is the “best evidence synthesis” (BES) review of the literature (Petticrew & Roberts, 2006). Essentially, BES entails first conducting a systematic quantitative review of the literature (e.g., Cooper, 1998; Cooper & Hedges, 1994; Ellis, 1991; D. Milne & James, 2000) and then applying a qualitative-constructivist methodology to the quantitative and qualitative review data (Polkinghorne, 1991; Strauss & Corbin, 1998). Specifically, BES entails a rigorous review of a carefully selected sample of empirical studies, employing a coding manual that defines the criteria to be used in extracting the quantitative and qualitative information, conducting reliability checks amongst the reviewers to ensure consistency and to reduce bias, and rigorous qualitative methods (Polkinghorne, 1991; Strauss & Corbin, 1998). The careful selection

and analysis of key studies provides a more focused and interpretable account of the extant literature (weak designs and ineffective interventions are excluded). Hence, BES utilizes both quantitative and qualitative research methodologies in the context of a literature review.

An example of BES is D. Milne and James (2000), who selected studies that reported successful outcomes of supervisor interventions, ones that were based on interpretable research designs (e.g.,  $N=1$  multiple baseline designs). Regarding the effectiveness of different methods of supervision, they found that a majority of the studies incorporated multiple instructional methods, including lectures and corrective feedback. The BES method permits relatively straightforward interpretations of the results, without the confounds of sample heterogeneity (i.e., considering a larger sample of studies with variable outcomes) and variable study rigor (i.e., combining poorly designed studies with more rigorous ones) faced by Ellis et al. (1996) and Ellis and Ladany (1997).

Using the BES method and following D. Milne and James's (2000) procedures, our intent was to extract the concepts used in a sample of successful and rigorous manipulations of clinical supervision in order to build a basic model of effective clinical supervision. Expressed as a research question, we wondered whether it was possible to explicate a conceptualization of supervision that was shared in common across the sampled studies. A general framework was utilized to organize these concepts, namely the idea of inter-linking contextual variables (moderators), supervision interventions (mediators), and outcomes (mechanisms of change; Baron & Kenny, 1986; Kraemer, Wilson, Fairburn, & Agras, 2004). The framework we selected was broad enough to be consistent with most theoretical approaches, reflected the inherent complexity of clinical supervision (Ellis & Ladany, 1997), and formed the basis for theory-driven evaluations (Chen, 1990).

## **OBJECTIVE**

In an attempt to address deficiencies in the clinical supervision literature, we aimed to build an inductively derived basic conceptualisation of effective, empirically supported supervision, one with appropriate complexity.

## METHOD

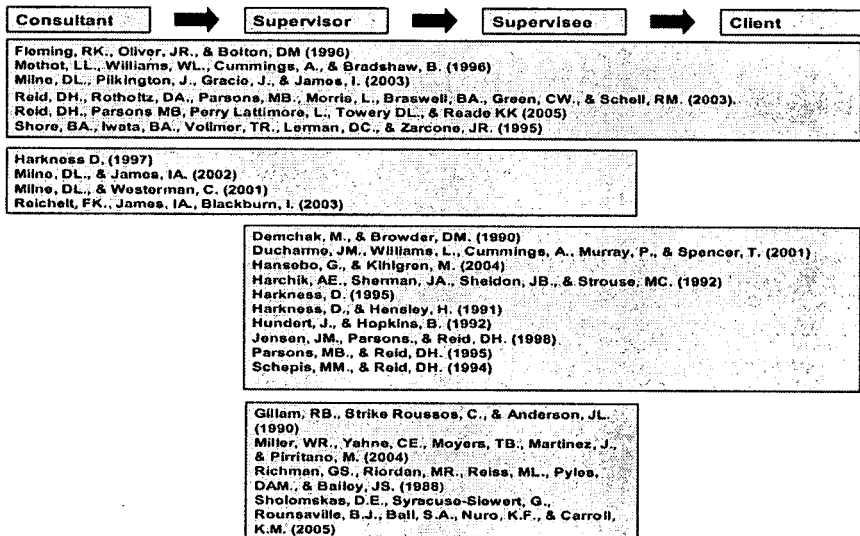
### *Inclusion Criteria and Coding Manual*

Drawing on current literature review methodology and guidelines (e.g., Cooper, 1998; Cooper & Hedges, 1994; Ellis, 1991; D. Milne & James, 2000), we explicated specific inclusion and exclusion criteria within a coding manual to obtain a sample of rigorous experimental studies. Studies from the clinical supervision literature were included in the present review if they satisfied eight criteria: (a) focused on clinical supervision and not training or mentoring; (b) was published in English, in the last 20 years; (c) was published in a peer-reviewed scientific journal; (d) included an interpretable and effective manipulation of supervision (e.g., used a rigorous research design to overcome problems of sample heterogeneity); (e) recorded the manipulation of supervision by direct observation; (f) recorded the behaviors of the supervisee; (g) was relevant to practice settings (i.e., had external validity); and (h) demonstrated clear benefits of clinical supervision (i.e., intended outcomes were achieved, thus was a "successful" study of supervision). There were no separate exclusion criteria (e.g., all client groups could potentially appear in the review sample). The manual also includes seven sections (e.g., research design; participants), detailing 32 criteria that address established dimensions of experimental research (e.g., the number and demographic characteristics of study participants). The manual provides the reviewer with guidelines on the coding of research studies against these 32 criteria. A copy is available on request to the first author.

### *Sample of Studies*

Over 100 potential studies were identified by searching electronic databases (IngentaConnect; ScienceDirect; Blackwell-Synergy; CSA Illumina; Ovid online; BIDS; Medline, and Streetwise). We also used Internet search engines (Google and MSN) to monitor promising current journals, browsed in libraries, searched the recent work of experts in the field, and studied the references cited within any of our selected papers. The authors also identified articles by examining studies listed in previous reviews (e.g., Ellis et al., 1996) and in books on supervision (e.g., Bernard & Goodyear, 2004). Of the over 100

FIGURE 1. The 24 Studies in the Systematic Review, Mapped onto the Educational Pyramid



potential studies for inclusion,  $N=24$  satisfied the inclusion/exclusion criteria. Figure 1 summarizes these studies, in relation to the four tiers of the educational pyramid (i.e., consultant/trainer, supervisor, therapist/supervisee, and patient). A representative example is the first study listed in Figure 1 (i.e., Fleming, Oliver, & Bolton, 1996). They evaluated a competency-based supervisor training program within two group homes for individuals with learning disabilities. The training entailed instructions, modeling, educational role-plays, and corrective feedback to the four participating supervisors. They were all graduates of human services-related areas, with variable experience as supervisors (between 8 months and 7 years), and an average age of 31 years. Within a multiple baseline design, observations were made of the frequency of nine supervisor behaviors (e.g., participative goal-setting, verbal instruction) and seven supervisee behaviors (e.g., prompting the client, providing a rewarding consequence). Clear effects were obtained for both the training and the supervision manipulations.

These 24 articles were mostly from the learning disabilities field, studied residential settings, and involved a range of professional groups. The effectiveness of supervision was typically evaluated in terms of multiple measures, including the reactions of the supervisees, indicators of their learning, and transfer to patients. Supervision was characterized by the use of feedback, support, and teaching. (A more detailed account of the supervision interventions and outcomes is available from the first author.)

### PROCEDURES

The manual was prepared to maximize the reliability and validity of the rating procedure. It included the construct definitions for *contextual variables* (moderators), *supervision interventions* (mediators or treatments), and *outcomes* (mechanisms). A contextual variable was defined as per a moderator variable; that is, a qualitative (i.e., nominal level) or quantitative variable that affected the direction and/or strength of the relationship between an independent and dependant variable (Baron & Kenny, 1986; Kraemer et al., 2002). That is, a moderator effect tests for whom and when a variable affects an outcome (Baron & Kenny, 1986; Frazier, Tix, & Barron, 2004; Kraemer et al., 2002). A mediator variable is an underlying change mechanism that explains the relation between the predictor and outcome variables and thereby tests how and why one variable affects an outcome (Baron & Kenny, 1986; Frazier et al., 2004; Kraemer et al., 2002). Hence, a supervision intervention was defined as a mediator variable, an event occurring during supervision (e.g., Socratic questioning of the supervisee) that had a main or interactive effect on the outcome variable (e.g., supervisee's improved understanding; see Kraemer et al., 2002). Finally, an outcome variable, or mechanism of change, was defined as the change process, the specification of one or more means through which the supervision intervention achieved an effect on the outcome variable (e.g., reflection increasing understanding). Thus, the supervision intervention was the independent variable and outcome was the dependent variable in the reviewed studies.

The raters were the first two authors. Interrater agreement/reliability was assessed at the outset of the present analysis for all 32 criteria within the coding manual and was found to be 75% exact agreement (based on the formula: sum of all agreements minus sum of all disagreements,

divided by sum of all agreements and disagreements, multiplied by 100 over 1). This was judged an acceptable level of interrater agreement, which was maintained at a "drift" assessment carried out between the raters for the final two studies (i.e., 82%). These data indicated that the 24 studies were coded consistently.

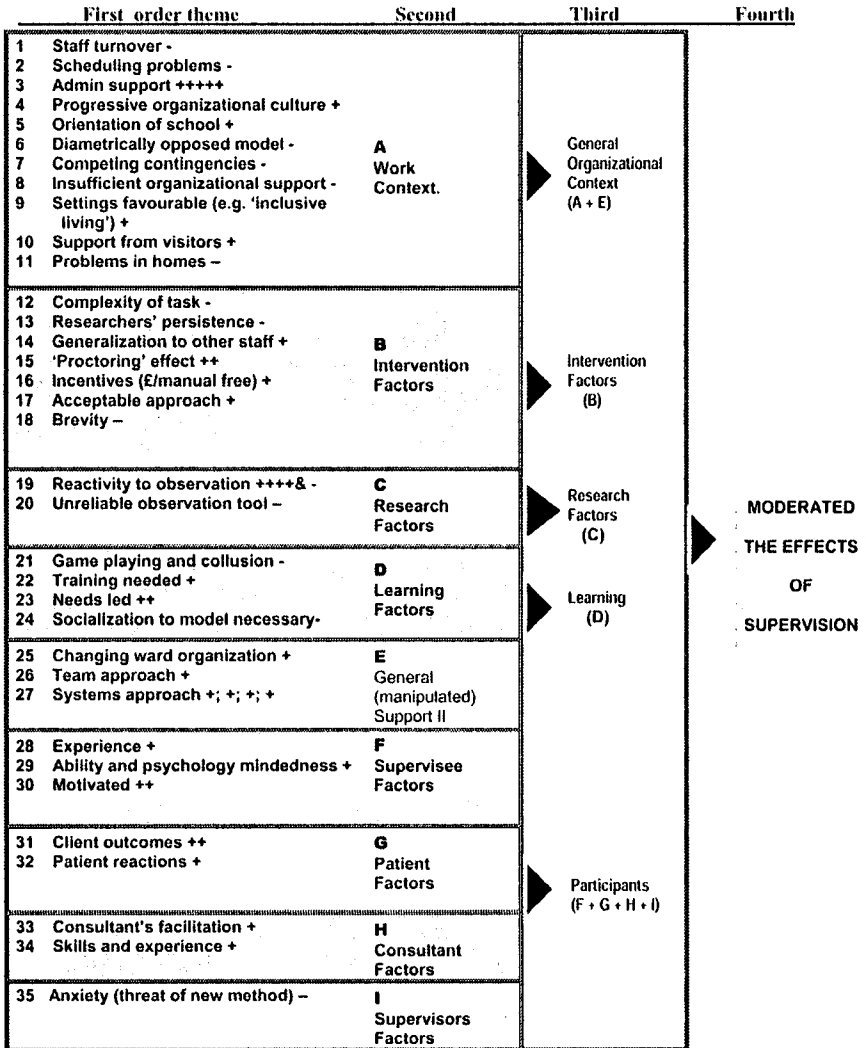
## RESULTS

### *Contextual Variables (Moderators)*

The authors of the 24 studies mentioned but did not manipulate 35 different contextual variables, such as staff turnover and administrative support. There were 50 instances of these 35 contextual variables, representing a mean of 2.1 contextual variables per study. Only two of the studies failed to identify a contextual variable (Gillam, Roussos, & Anderson, 1990; Shore, Iwata, Vollmer, Lerman, & Zarcone, 1995). According to the authors of the other 22 studies, the contextual variables mostly enhanced the obtained results (in 34 of the 50 instances; i.e., 68% of contextual variables were judged as facilitative). This suggests that the reviewed studies used appropriately complex analyses and were conducted in direct service settings and conditions. Figure 2 is a dendrogram, providing a summary of the contextual variable information in the reviewed 24 studies. The most frequently mentioned contextual variable was administrative support (Item 3 in Figure 2). All five references to administrative support referred to it as a positive influence on the supervision intervention (+symbol in Figure 2). Using inductive qualitative and consensus validation procedures, the 35 contextual variables were categorized by the authors into five themes to distill the fundamental latent concepts. The five contextual themes were the general organizational context, the participants, intervention factors, research influences, and learning. As indicated in Figure 2, each of these themes had a combination of positive and negative influences on the effectiveness of supervision. To illustrate these influences for the organization context theme, the amount and type of administrative support and the process by which the support was provided may predict how a supervisor behaves (Hundert & Hopkins, 1992). The effect can also be negative: "staff turnover influenced this study" (Demchak & Browder, 1990, p. 161).



FIGURE 2. A Summary of Moderators Identified by the Authors



In summary, nearly all of the studies included within this review listed and interpreted the valence of contextual variables. Although not manipulated within their studies, we judged these factors to have had a generally positive influence on their findings.

*Supervision Interventions (Mediators)*

Twenty-six different supervision interventions (i.e., mediators or treatments) were reported within the 24 reviewed studies (see Table 1). The most frequently cited of these was teaching and instruction (75% of studies), followed by the use of corrective

TABLE 1. Supervision Interventions in the 24 Reviewed Studies, Ranked by Frequency of Occurrence

|  | No. Studies | % Studies Prevalence | Rank |
|--|-------------|----------------------|------|
| Training (teaching skills/instruction)   | 18          | 75                   | 1    |
| Feedback   | 15          | 63                   | 2    |
| Observing (live or recorded)   | 10          | 42                   | 3    |
| Goal-setting   | 9           | 38                   | 4.5  |
| Question-and-answer (information gathering; clarifying)  | 9           | 38                   | 4.5  |
| Modeling (demonstration, live or video/audio)  | 7           | 29                   | 6    |
| Planning (including managing; agenda; next meeting)  | 6           | 25                   | 7.5  |
| Reinforcement/praise/support   | 6           | 25                   | 7.5  |
| Discussion (review)  | 5           | 21                   | 10   |
| Prompts (verbal and written reminders, handout, etc.)  | 5           | 21                   | 10   |
| Role-play  | 5           | 21                   | 10   |
| Explanation (rationale provided socialization to model)  | 4           | 17                   | 12.5 |
| Monitoring (evaluating) client benefit/supervisee performance  | 4           | 17                   | 12.5 |
| Review/reflection  | 5           | 21                   | 12.5 |
| Summarizing  | 4           | 17                   | 14.5 |
| Challenging (rethink)  | 3           | 13                   | 15   |
| Self-monitoring  | 1           | 4                    | 18   |
| Listening  | 2           | 8                    | 18   |
| Problem-solving  | 2           | 8                    | 18   |
| Rehearsal of skills  | 2           | 8                    | 18   |
| Self-disclosure  | 3           | 13                   | 18   |
| Collaborating (working together)   | 1           | 4                    | 23   |
| Confidence building (efficacy)   | 1           | 4                    | 23   |
| Disagreeing  | 1           | 4                    | 23   |
| Formulating (modeling problem)   | 1           | 4                    | 23   |
| Understanding checked multiple/varied elements counted only once (e.g., different types of training) | 1           | 4                    | 23   |

feedback, which was reported in 15 of the studies (63%). Live or video-based observation of the supervisee (42% of studies), then goal-setting (38%), and question-and-answer methods (including information-gathering and clarifying: 38%) were the next most frequently cited interventions. The total number of supervision interventions used within the 24 studies was 130, signifying that each study used multiple interventions ( $M=5.4$ ,  $SD=4.32$ ). For example, Miller, Yahne, Moyers, Martinez, and Pirritano (2004) used an initial training workshop lasting 2 days, collaborative problem-solving, the use of role-play rehearsal, and modeling to develop the supervisees' competence, and the provision of positive reinforcement.

In summary, although not as numerous as the identified contextual variables, 26 supervision interventions (mediators) were identified within the 24 empirical studies, and each of the reviewed studies tended to use over 5 of these methods. The interventions are consistent with current theories and models of good practice for the development of the supervisee (Kaslow et al., 2004), as is the use of multiple interventions (e.g., Bransford, Brown, & Cocking, 2000).

### *Outcomes (Mechanisms of Change)*

The authors of the 24 studies defined 28 different outcomes (mechanisms of change). These included changes to the supervisees' attitudes, increased emotional self-awareness, changes in supervisees' motivation, and improved skills. To facilitate understanding of these supervision outcomes, we used Kolb's (1984) experiential learning cycle. Kolb identified four modes of learning from experience (experiencing, reflecting, conceptualizing, and experimenting). Following the inductive qualitative and consensus-validation procedure once more, we classified 23 of the 28 outcomes (82%) according to Kolb's experiential learning cycle (see Table 2). Five examples of outcomes were classified as "other" (including general references to the supervisees learning and to self-monitoring).

Table 2 indicates that the most frequently cited outcomes (mechanism of change) was "experiencing," which was reported in 12 of the studies (43%). Hansebo and Kihlgren (2004) used reflection within supervision to increase the supervisees' awareness and knowledge about their influence on the quality of patient care. That is, the supervision intervention changed the supervisees' attitudes and

TABLE 2. Supervision Outcomes Reported in 24 Reviewed Studies, Ranked by Frequency of Occurrence

|   | No. Studies<br>Specifying | % Studies | Rank |
|---|---------------------------|-----------|------|
| Experiencing (attitude change; affective awareness; motivation/reinforcement) | 12                        | 50        | 1    |
| Other (general "learning"; self-monitoring/supervision)                       | 5                         | 20        | 2    |
| Reflection (to raise self-awareness/evaluation, positive and negative)        | 3                         | 12        | 3    |
| Planning (increased attention to goals/focusing)                              | 4                         | 16        | 4    |
| Conceptualizing   | 2                         | 8         | =5.5 |
| Experimenting ("exposure"/learning by doing)                                  | 2                         | 8         | =5.5 |

perspectives toward their patients, viewing the patient as a "unique person—with resources and capabilities" (Hansebo & Kihlgren, 2004, p. 273).

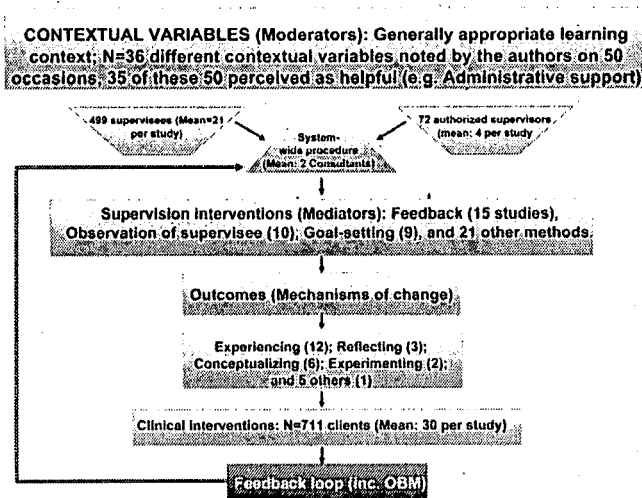
In summary, although the reviewed authors identified a large number of outcome variables, 82% of the reviewed studies described outcomes consistent with the experiential learning cycle (Kolb, 1984). Given that experiential methods are "powerful" and "essential" (Kaslow et al., 2004, p. 780), Kolb's (1984) model merits empirical attention in supervision.

### *A Basic Model of Clinical Supervision*

Based on the above analyses, we delineated a basic model of clinical supervision (see Figure 3). We summarized the main contextual variables, which generally provided a facilitating learning context for the other variables in the model. Figure 3 also summarized the different supervision interventions that were manipulated by the supervisors and that favorably influenced the therapist supervisee's clinical interventions provided to 711 patients (i.e., a mean of 30 patients per study, or approximately 1.2 patients per supervisee). Finally, the authors of the 24 reviewed studies described various corrective feedback systems to revise and improve the supervision interventions.

Illustrating the supervision model in practice, Reid et al. (2003) developed and evaluated a state-wide approach to use "positive

FIGURE 3. The Basic Model of Clinical Supervision



behavior support" (PBS) for people with learning disabilities in South Carolina. The contextual variables were a trend to create more inclusive community living arrangements state-wide. Supervision interventions included observing the use of the PBS behaviors by the supervisor until supervisees mastered implementing PBS with their patients. Reid et al. (2003) regarded the supervisees' satisfaction with the content and process of the approach, together with their mastery and motivation, to be the outcome (classified as the experiencing mode in Table 2). As an example of a corrective feedback loop, Reid et al. monitored the effectiveness and generalisation of the PBS activities over time and setting, so that they could adjust the supervision and ensure successful transfer to the benefit of the patients. Participants also rated the supervision interventions, such as its usefulness, any benefits, and key contextual variables.

## DISCUSSION

Our main objectives were to conduct a best evidence synthesis review (Petticrew & Roberts, 2006) of 24 rigorous studies that had successfully manipulated supervision and to codify a basic model of

clinical supervision based on the review data. Although few of the reviewed studies explicitly articulated a conceptual model, there were clear references made to contextual variables (moderators), supervision interventions (mediators), and treatment outcomes (mechanisms of change). These data were the basis for an inductive crystallization of a basic model, our second objective. A dendrogram of the 32 contextual variables (Figure 2) indicated that five themes captured the organizational systems within which the reviewed studies took place. Of note, the organizational contextual influences included staff turnover and administrative support. In turn, these contextual variables were linked to 26 supervision interventions, which collapsed into the categories of teaching (75% of studies), providing corrective feedback (63%), observing (42%), and other popular supervision interventions (see Table 1). Next, Table 2 presents the outcomes, the processes that might explain how such supervision interventions developed competence and clinical effectiveness in the supervisees. Eighty-two percent of the 28 different outcomes articulated in the 24 studies were categorized into one of the four modes of the experiential learning cycle (Kolb, 1984). This result suggested a central role for experiential learning within these studies.

### *A Consensus?*

Our third and final objective was to ask how well this basic model agreed with the broader clinical supervision literature. Our findings were consistent with the results from other reviewers (e.g., Gregoire, Propp, & Poertner, 1998; Harkness & Poertner, 1989; D. Milne & James, 2000). In terms of contextual variables, we found that supervision was indeed "embedded in complex social networks" (Falender et al., 2004, p. 779); the work environment was perhaps the most critical yet least controllable influence on supervision (Gregoire et al., 1998); supervision was a training process in which multiple forms of supervision occurred (Harkness & Poertner, 1989); and the importance of measuring clinical outcomes of supervision was reaffirmed (Holloway & Neufeldt, 1995; Stein & Lambert, 1995).

Like previous reviewers, we found that researchers employed more than one supervision method (D. Milne & James, 2000), utilized in a systematic supervisee development package that included teaching, modeling, rehearsal, and corrective feedback (e.g., Falendar et al.,

2004; Kilminster & Joley, 2000; Watkins, 1997), and that feedback and specific skills training were the most commonly used supervision intervention (D. Milne & James, 2000; Norcross & Halgin, 1997). Norcross and Halgin (1997) underscored the need to use a blend of supervision methods, including a wide range of techniques. This might involve "didactic presentations, reading assignments, open-ended discussions, personal modelling, experiential activities, video demonstrations, case examples and mini-case conferences" (p. 209).

As for the outcomes within supervision, reflection was perhaps the most frequently cited example within the literature (Kilminster & Joley, 2000). Corroboration of the remaining modes of experiencing, conceptualizing, and planning can be found within the *Handbook of Psychotherapy Supervision* (Watkins, 1997).

### *Critique*

A criticism of the present study was that, like much of the research in supervision, it lacked sufficient rigor. Specifically, the categories (e.g., "organizational climate"; "feedback") used to capture the contextual variables, supervision interventions, and outcomes were diverse and often poorly defined in the original articles, making a significant degree of inference necessary. In addition, determining the extent to which authors were indicating the salience of these variables within their studies was at times problematic. This was because few studies set out to test explicitly stated models, so we regularly inferred the conceptualization underpinning the reviewed studies. This inevitably introduced difficulties. However, against this threat the authors produced a manual and used a consensus validation method and demonstrated that we could code studies in a consistent fashion. Nonetheless, given the inferential nature of the task, the possibility of bias arising from the inclusion of three of the first authors own coauthored studies, and the ambiguity of much of the research articles, it would be appropriate to treat the integrative model (Figure 3) as tentative and awaiting empirical testing. A further issue is that, having reviewed a preponderance of learning disabilities studies, the generalizability of the derived model cannot be assumed, either to other theoretical orientations (the present studies are broadly behavioral) or to other clinical specializations. For example, it says little about the attributes of the participating supervisors and supervisees or about the supervisory alliance. Nonetheless, hopefully its basic

nature will enable it to serve as a reference point to stimulate future research.

## CONCLUSIONS

Our basic clinical supervision model is unusual in being evidence based, having been generated inductively from a highly selected sample of 24 successful studies of clinical supervision. Consistent with recent reviews, the model proposes that supervision is a complex activity, one that is contextualized by at least five major types of contextual variables and implemented (mediated) through over two dozen supervision interventions. These interventions are thought to work primarily by promoting experiential learning (i.e., the outcome or mechanism of change). At least within the current sample, there appeared to be positive clinical outcomes from this evidence-based model of supervision. The basic model had good conceptual rigor, as it offered improved clarity about what supervision is and how it works under naturalistic (i.e., complex) conditions. This conceptualization therefore improves on other general models that lack sufficient specificity to be tested (i.e., they may not be scientific models; e.g., Kaufman & Schwartz, 2003). By contrast, our basic model enables theoretically grounded hypotheses to be developed and stated (i.e., the model has "hypothesis validity": Wampold, Davis, & Good, 1990). For instance, the model suggests (through its specification of the outcomes) how the supervisor facilitates learning in the supervisee (see Table 1), a weakness in prior model-building efforts (Watkins, 1995b). Next, research that operationalizes and tests the hypotheses that follow from this basic model is required, perhaps initially through experimental  $N=1$  studies (e.g., Dennin & Ellis, 2003) that examine the specific relationships between these outcomes and the relevant learning and clinical outcomes. In this way, we can give supervision the greater importance it merits (Watkins, 1997), from an improved conceptual foundation.

## ACKNOWLEDGEMENT

We are grateful to Karen Clark for help in word processing this manuscript, to Judy Preece for help in preparing the figures, and to



other colleagues for their help with the general supervision research program of which this is a part (especially Chris Dunkerley, Sarah Wharton, Ian James, and Liz Kennedy).

## REFERENCES

(\*indicates reviewed article)

- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182.
- Beinart, H. 2004. Models of supervision and the supervisory relationship and their evidence-base. In I. Fleming & L. Steen (Eds.), *Supervision and clinical psychology* (pp. 36-50). New York: Brunner Routledge.
- Bernard, J. M. (1979). Supervisory training: A discrimination model. *Counselor Education and Supervision, 19*, 60-68.
- Bernard, J. M., & Goodyear, R. K. (2004). *Fundamentals of clinical supervision*. (3rd ed.). Boston, MA: Pearson Education.
- Borders, L. D. (1989). A pragmatic agenda for developmental supervision research. *Counselor Education and Supervision, 29*, 16-24.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn*. Washington, DC: National Academy Press.
- Chen, H. T. (1990). *Theory driven evaluations*. London: Sage.
- Colquitt, J. A., LePine, J. A., & Noe, R. A. (2000). Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research. *Journal of Applied Psychology, 85*, 678-707.
- Cooper, H. M. (1998). *Integrating research: A guide for literature reviews*. (3rd ed.). Newbury Park, CA: Sage.
- Cooper, H. M., & Hedges, L. V. (Eds.). (1994). *Handbook of research synthesis*. New York: Russell Sage Foundation.
- \*Demchak, M., & Browder, D. M. (1990). An evaluation of the pyramid model of staff training in group homes for adults with severe handicaps. *Education and Training in Mental Retardation, 25*, 150-163.
- Dennin, M. K., & Ellis, M. V. (2003). Effects of a method of self-supervision for counselor trainees. *Journal of Counseling Psychology, 51*, 69-83.
- Department of Health. (2004). *Organising and delivering psychological therapies*. London: Author.
- Department of Health. (2005). *Delivering the NHS improvement plan: The workforce contribution*. London: Author.
- Doss, B. D., Thum, Y. M., Sevier, M., Atkins, D. C., & Christenson, A. (2005). Improving relationships: Mechanisms of change in couple therapy. *Journal of Consulting and Clinical Psychology, 73*, 624-633.
- Duchamne, J. M., Williams, L., Cummings, A., Murray, P., & Spencer, T. (2001). General case quasi-pyramidal staff training to promote generalization of

- teaching skills in supervisory and direct-care staff. *Behavior Modification*, 25, 233-254.
- Ellis, M. V. (1991). Conducting and reporting integrative research reviews: Accumulating scientific knowledge. *Counselor Education and Supervision*, 30, 225-237.
- Ellis, M. V., & Ladany, N. (1997). Inferences concerning supervisees and clients in clinical supervision: An integrative review. In C. E. Watkins, Jr. (Ed.), *Handbook of psychotherapy supervision* (pp. 447-507). Chichester: Wiley.
- Ellis, M. V., Ladany, N., Krenal, M., & Schult, D. (1996). Clinical supervision research from 1981 to 1993: A methodological critique. *Journal of Counseling Psychology*, 43, 35-50.
- Falender, C., & Shafranske, E. (2004). *Clinical supervision: A competence-based approach*. Washington, DC: American Psychological Association.
- Falender, C. A., Cornish, J. A. E., Goodyear, R., Hatcer, R., Caslow, N. J., Levental, G., et al. (2004). Defining competencies in psychology supervision: A consensus statement. *Journal of Clinical Psychology*, 60, 771-785.
- \*Fleming, R. K., Oliver, J. R., & Bolton, D. M. (1996). Training supervisors to train staff: A case study in a human service organization. *Journal of Organizational Behaviour Management*, 16, 3-25.
- Frazier, P. A., Tix, A. P., & Barron, K. E. (2004). Testing moderator and mediator effects in counseling psychology research. *Journal of Counseling Psychology*, 51, 115-134.
- \*Gillam, R. B., Roussos, C., & Anderson, J. L. (1990). Facilitating changes in supervisees' clinical behaviours: An experimental investigation of supervisory effectiveness. *Journal of Speech and Hearing Disorders*, 55, 729-739.
- Gregoire, T. K., Propp, J., & Poertner, J. (1998). The supervisor's role in the transfer of training. *Administration in Social Work*, 22, 1-18.
- \*Hansebo, G., & Kihlgren, M. (2004). Nursing home care: Changes after supervisions. *Journal of Advanced Nursing*, 45, 269-279.
- \*Harchik, A. E., Sherman, J. A., Sherlodon, J. D., & Strouse, M. C. (1992). Ongoing consultation as a method of improving performance of staff members in a group home. *Journal of Applied Behavior Analysis*, 25, 599-610.
- \*Harkness, D. (1995). The art of helping in supervised practice: Skills, relationships and outcomes. *Clinical Supervision*, 13, 63-76.
- \*Harkness, D. (1997). Testing interactional Social Work theory: A panel analysis of supervised practice and outcomes. *The Clinical Supervisor*, 15, 33-50.
- \*Harkness, D., & Hensley, H. (1991). Changing the focus of social work supervision: Effects on client satisfaction and generalised contentment. *Social Work*, 36, 506-512.
- Harkness, D., & Poertner, J. (1989). Research and social work supervision: A conceptual review. *Social Work*, 34, 115-119.
- Holloway, E. L. (1995). *Clinical supervision: A systems approach*. Thousand Oaks, CA: Sage.
- Holloway, E. L., & Neufeldt, S. A. (1995). Supervision: Its contribution to treatment efficacy. *Journal of Consulting and Clinical Psychology*, 63, 207-213.

- \*Hundert, J., & Hopkins, B. (1992). Training supervisors in a collaborative team approach to promote peer interaction of children with disabilities in integrated pre-schools. *Journal of Applied Behavior Analysis*, 25, 385-400.
- \*Jenson, J. M., Parsons, M. B., & Read, D. H. (1998). Supervisory training for teachers: Multiple, long term effects in an education programme for adults with severe learning disabilities. *Research in Developmental Disabilities*, 19, 449-463.
- Kaslow, F. W. (1986). *Supervision and training: Models, dilemmas, and challenges*. London: Haworth Press.
- Kaslow, N. J., Borden, K. A., Collins, F. L., Forrest, L., Illfelder-Kaye, J., Nelson, P. D., et al. (2004). Competencies conference: Future directions in education and credentialing in professional psychology. *Journal of Clinical Psychology*, 60, 699-712.
- Kaufman, J., & Schwartz, T. (2003). Models of supervision: Shaping professional identity. *The Clinical Supervisor*, 22, 143-158.
- Kerlinger, F. N., & Lee, H. B. (2000). *Foundations of behavioral research* (4th ed.). New York: Harcourt.
- Kilminster, S. M., & Joley, B. C. (2000). Effective supervision in clinical practice settings: A literature review. *Medical Education*, 34, 827-840.
- Kolb, D. A. (1984). *Experiential learning*. Upper Saddle River, NJ: Prentice-Hall.
- Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in RCT-S. *Archives of General Psychiatry*, 59, 877-883.
- Leary, D. E. (1990). *Metaphors in the history of psychology*. Cambridge: University Press.
- Leddick, G. R., & Bernard, J. M. (1980). The history of supervision: A critical review. *Counselor Education and Supervision*, 20, 186-196.
- \*Methot, L. L., Williams, W. L., Cummings, A., & Bradshaw, B. (1996). Measuring the effects of a manager's supervisor training programme through the generalised performance of managers, supervisors, front line staff and clients in a human service setting. *Journal of Organizational Behavioural Management*, 16, 3-34.
- \*Miller, W. R., Yahne, C. E., Moyers, T. B., Martinez, J., & Pirritano, M. (2004). A randomized trial of methods to help clinicians learn motivational interviewing. *Journal of Consulting and Clinical Psychology*, 72, 1050-1062.
- \*Milne, D., & James, I. A. (2000). A systematic review of effective cognitive behavioural supervision. *British Journal of Clinical Psychology*, 39, 111-127.
- Milne, D., & James, I. A. (2005). Clinical supervision: Ten tests of the tandem model. *Clinical Psychology Forum*, 151, 6-9.
- Milne, D. L., & James, I. A. (2002). The observed impact of training on competence in clinical supervision. *British Journal of Clinical Psychology*, 41, 55-72.
- \*Milne, D. L., Pilkington, J., Gracie, J., & James, I. A. (2003). Transferring skills from supervision to therapy: a qualitative and quantitative  $N=1$  analysis. *Behavioural and Cognitive Psychotherapy*, 31, 193-202.
- \*Milne, D. L., & Westerman, C. (2001). Evidence-based clinical supervision: Rationale and illustration. *Clinical Psychology and Psychotherapy*, 8, 444-457.

- Moyers, T. B., Miller, W. R., & Hendrickson, S. M. L. (2005). How does motivational interviewing work? Therapist inter-personal skill predicts client involvement within motivational interviewing sessions. *Journal of Consulting and Clinical Psychology, 73*, 590–598.
- National Health Service, Centre for Reviews and Dissemination. (2001). Undertaking systematic reviews of research on effectiveness. Retrieved February 1 2004 from <http://www.york.ac.uk/inst/crd/report4.htm>
- Norcross, J. C., & Halgin, R. P. (1997). Integrative approaches to psychotherapy supervision. In C. E. Watkins, Jr. (Ed.), *Handbook of psychotherapy supervision*. Chichester: Wiley.
- \*Parsons, M. B., & Reid, D. H. (1995). Training residential supervisors to provide feedback from maintaining staff teaching skills with people who have severe disabilities. *Journal of Applied Behavior Analysis, 28*, 317–322.
- Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences*. Oxford: Blackwell Publishing.
- Polkinghorne, D. E. 1991. Qualitative procedures for counseling research. In C. E. Watkins, Jr., & L. J. Schneider (Eds.), *Research in counseling* (pp. 163–204). Hillsdale, NJ: Erlbaum.
- Popper, K. R. (1972). *Conjectures and refutations: The growth of scientific knowledge*. London: Routledge.
- \*Reichelt, F. K., James, I. A., & Blackburn, I. M. (2003). Impact of training on rating competence in cognitive therapy. *Journal of Behaviour Therapy and Experimental Psychiatry, 34*, 87–99.
- \*Reid, D. H., Parsons, M. B., Lattimore, P. L., Towery, D. L., & Reade, K. K. (2005). Improving staff performance through clinical application of outcome management. *Research in Developmental Disabilities, 26*, 101–116.
- \*Reid, D. H., Rotholtz, D. A., Parsons, M. B., Morris, L., Braswell, B. A., Green, C. W., et al. (2003). Training human service supervisors in aspects of PBS. *Journal of Positive Behaviour Interventions, 5*, 35–46.
- \*Richman, G. S., Riordan, M. R., Reiss, M. L., Pyles, D. A. M., & Bailey, J. S. (1988). The effects of self monitoring and supervisor feedback on staff performance in a residential setting. *Journal of Applied Behavioural Analysis, 21*, 401–409.
- Rodenhauser, P. (1997). Psychotherapy supervision: Pre-requisites and problems in the process. In C. E. Watkins, Jr. (Ed.), *Handbook of psychotherapy supervision* (pp. 527–548). New York: Wiley.
- \*Schepis, M. M., & Reid, D. H. (1994). Training direct service staff in congregate settings to interact with people with sever disabilities: A quick, effective and acceptable programme. *Behavioural Interventions, 9*, 13–26.
- Serlin, R. C. (1987). Hypothesis testing, theory building, and the philosophy of science. *Journal of Counseling Psychology, 34*, 365–371.
- \*Sholomskas, D. E., Syracuse-Siewert, G., Rounsaville, B. J., Ball, S. A., Nuro, K. F., & Carroll, K. M. (2005). We don't train in vain: A dissemination trial of three strategies of training clinicians in CBT. *Journal of Consulting and Clinical Psychology, 73*, 106–115.

- \*Shore, B. A., Iwata, B. A., Vollmer, T. R., Lerman, D. C., & Zarcone, J. R. (1995). Pyramidal staff training in the extension of treatment for severe behaviour disorders. *Journal of Applied Behaviour Analysis*, 28, 323-332.
- Stein, D. M., & Lambert, M. J. (1995). Graduate training in psychotherapy: Are therapy outcomes enhanced?. *Journal of Consulting and Clinical Psychology*, 63, 182-196.
- Stoltenberg, C. D., & Delworth, U. (1987). *Supervising counsellors and therapists: A developmental approach*. San Francisco: Jossey-Bass.
- Stoltenberg, C. D., & McNeil, B. W. (1997). Clinical supervision from a developmental perspective: Research and practice. In C. E. Watkins, Jr. (Ed.), *Handbook of psychotherapy supervision* (pp. 184-202). Chichester: Wiley.
- Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. (2nd ed.). Thousand Oaks, CA: Sage.
- Tharenou, P. (2001). The relationship of training motivation to participation in training and development. *Journal of Occupational and Organizational Psychology*, 74, 599-621.
- Tracey, T. J. G., & Glidden-Tracey, C. E. (1999). Integration of theory, research design, measurement, and statistics: Toward a reasoned argument. *The Counseling Psychologist*, 27, 299-324.
- Wampold, B. E., Davis, B., & Good, R. H. (1990). Hypothesis validity in clinical research. *Journal of Consulting and Clinical Psychology*, 58, 360-367.
- Wampold, B. E., & Holloway, E. L. (1977). Methodology, design and evaluation in psychotherapy supervision research. In C. E. Watkins, Jr. (Ed.), *Handbook of psychotherapy supervision* (pp. 11-30). Chichester: Wiley.
- Warr, P. B. 1980. An introduction to models in psychological research. In A. J. Chapman & D. M. Jones (Eds.), *Models of man*. Leicester: BPS.
- Watkins, C. E., Jr. (1995a). Psychotherapy supervisor development: On musings, models, and metaphor. *The Journal of Psychotherapy Practice and Research*, 4, 150-158.
- Watkins, C. E., Jr. (1995b). Psychotherapy supervisor and supervisee: Developmental models and research nine years later. *Clinical Psychology Review*, 15, 647-680.
- Watkins, C. E., Jr. (1997). *Handbook of psychotherapy supervision*. Chichester: Wiley.
- Worthington, E. L. (1987). Changes in supervision as counselors and supervisors gain experience: A review. *Professional Psychology: Research and Practice*, 18, 189-208.