Is Supervisor Training Effective? A Pilot Investigation of Clinical Supervisor Training Program

Abstract

Objective: Although competency-based training of supervisors is now compulsory in many countries, there has been limited evidenced-based literature to guide the profession as to the most effective ways of training supervisors. The aim of the current pilot study was to examine the effects of a supervisor training program on supervisor, supervisee, and evaluator perceptions of supervisory competence.

Method: Ten female and two male supervisors employed by a state-wide counselling service participated in the study. Data, including taped supervision sessions relating to supervisor, supervisee, and evaluator perceptions of supervisory competence, skill, and process were collected at three time points; two prior to a supervisor training program and one following the program.

Results: The evaluator observed significant improvements in supervision competence from pre to post training. Although supervisors reported no significant change in competence as a result of training, they did report areas of positive changes in supervisory practice. Across time supervisors also reported a significant decrease in their perception of the supervisor-supervisee alliance on one measure, but this was not noted by the evaluator or supervisee, or on other measures of supervisory alliance.

Conclusions: Evaluator ratings of supervisor competency and qualitative supervisor feedback provide preliminary support for the effectiveness of supervisor training. However, the pilot study was underpowered and some of the measures require further psychometric testing, which will need to be addressed in future research.

Keywords: Supervision, supervisor competencies, supervisor effectiveness, supervisor training, clinical supervision

A Pilot Investigation of Supervisor Training on Supervision Competencies

The need for psychology practitioners to undergo supervision competency training has generally become an accepted requirement across countries. Despite the strong support for supervision training, there is limited evidence-based literature to guide the psychology profession on effective ways of training supervisors (Kavanagh et al., 2008; Spence, Wilson, Kavanagh, Strong, & Worrall, 2001; Whitman, Ryan, & Rubenstein, 2001). Some factors limiting the confidence in the evidence base for supervisor training include differing views of what supervision training needs to be effective at changing (Milne & Reiser, 2014), different approaches to supervision training, studies of supervision effectiveness with different professions, small sample sizes, reliance on self-report, and the use of psychometrically weak instruments (Milne, Sheikh, Pattison, & Wilkinson, 2011; Watkins 2012). This current paper attempts to advance our knowledge of the effectiveness of supervision training by addressing some of these limitations. A unique aspect of the study is its focus on the benefits of supervisor training for already practising and experienced supervisors, a group previously not studied, but who are increasingly undertaking supervisor training programs in order to meet regulatory requirements.

Two important reviews of supervisor training effectiveness have recently been conducted (Milne, et al., 2011; Watkins, 2012). Milne et al. (2011) conducted a systematic review of 11 controlled studies, and concluded that there was preliminary support for supervision training in developing supervisor skill acquisition and behavioural change, knowledge; generalisation of skills; and supervisee satisfaction. The top three specific elements of supervision training which appeared to help achieve these outcomes were: 1) corrective feedback; 2) educational role-play; and 3) modelling/observational learning.

Although the Milne et al. (2011) review was positive and gave helpful conclusions to guide supervision training, the majority of the studies in the review were based on training of

teachers, paraprofessionals, and human services staff, with only two of the eleven studies specific to counselling or psychology (i.e. Bambling, King, Raue, Schweitzer & Lambert, 2006; McMahon & Simons, 2004). Of these, one examined effects on client outcome, the other on supervisor competencies. Bambling et al. (2006) found that both working alliance based supervision training models increased client report of working alliance and reduced symptoms of depression. However, a limitation of the study was it did not also measure the effects of supervision training on supervisory alliance or competencies, and some argue that supervision training research needs to focus on the effects of training on competencies (Reiser & Milne, 2014). The other supervisor training program for counsellors/ psychotherapists reviewed by Milne et al. (2011) evaluated the effectiveness of a 4-day supervisor workshop on supervisors' self-reported knowledge, confidence, and skills (McMahon & Simons, 2004). The study was the only study to use a longitudinal experimental design. Supervisors in the experimental group improved significantly on supervision competencies relative to control group of supervisors. However, supervisor knowledge was measured using a closed format self-report measure and future research is needed that measures supervisor skills and behaviour from the perspective of the supervisee.

The reliance on supervisor self-report - in the absence of reports from supervisee or an objective external evaluator - is a major criticism made by Watkins (2012), author of the second major review in the area. Of the 20 studies included in the Watkins (2012) review – all of which reported on supervision training specific to psychotherapy and focused on intervention effects on supervisor competencies – all but two (Milne & Westerman, 2001; Milne & James, 2002) relied on self-report to evaluate training effectiveness. Milne and Westerman (2001) video-taped and coded the supervision sessions of two supervisors and noted that supervision training consisting of 1- hour feedback and consultancy which led to a

modest increase in the number of desirable supervisor behaviours (e.g. checking theoretical knowledge and guiding experiential learning).

In a subsequent study, Milne and James (2002) found that supervision consultancy increased supervisor effectiveness (as rated by observers) on guiding experiential learning, which co-occurred with an increase in supervisee conceptualising skills, use of role play, and expressing emotion or attitude to a situation. The two Milne and colleagues' studies have a number of important strengths, including: experiential learning through view and review of video-taped supervision session; use of an objective behavioural rating scale; and a longitudinal analysis. However, the single case design and supervision consultancy approach which does not represent the norm of supervisor training, limit the generalisability.

The current study:

The current study advances the field by investigating the change in supervisor self-report, supervisee report, and an external evaluator report of supervisor competencies, supervisory alliance, supervisor confidence, and supervisor satisfaction from before a supervision training workshop (two baseline assessments were collected to test if variables stayed stable across time without training) and shortly after the training workshop.

The supervision training program was similar to those most commonly applied in international training programs (i.e. a didactic workshop combined with experiential learning gained through feedback on video-taped sessions of supervision) and included the three elements that the Milne et al. (2011) review identified as common among supervision training programs (namely corrective feedback, educational role-play, and modelling/observational learning). It was predicted that (1) supervisory competence would be stable over the baseline period (between time 1 and 2), and (2) improve over the intervention period (between time 2 and 3) as reported by multiple informants (i.e., supervisor, supervisee and evaluator). It was

also predicted that (3) supervisors would report increased confidence and satisfaction with supervision over time, and (4) positively evaluate the training program.

Methods

Participants

All of the 18 supervisors employed to provide supervision for practitioners in a state-wide counselling service were invited to participate in the study and 12 agreed (10 females; 2 males). Ages ranged from 38 to 65 (M = 52.08, SD = 8.66) years. The supervisors were psychologists, counsellors, or mediators by profession. They reported providing an average of 10.59 (SD = 3.22) hours of supervision per week.

Materials

Supervisor, supervisee, and evaluator report questionnaires were used to examine change in supervisory competence, alliance, and confidence. A number of widely used measures were used: the Supervisory Working Alliance Inventory (SWAI) by Efstation, Patton and Kardash (1990); the Leeds Alliance in Supervision Scale (LASS) by Wainwright (2010); the Supervision Rating Scale (SRS) by Duncan & colleagues (2003); and the Supervisee Satisfaction Questionnaire (SSQ) by Ladany, Hill, Corbett and Nutt (1996). Because of their wide use and acceptance in the field, these questionnaires are not described here. Other less common measures were also utilised.

Evaluation of Recorded Supervision Session (ERS). The ERS is an original questionnaire designed by the first author to assess supervisory competence in relation to a recorded supervision session. The ERS is completed by both supervisor and evaluator, and is designed to capture supervisory competence across seven domains; positive supervisor characteristics, a range of roles as appropriate to the session, effective administration and structure in the session, effective supervisory skills, evidence of an effective supervisory alliance, degree of knowledge of content specific to speciality area, and appropriate ethical

behaviour and knowledge. Supervisors and evaluators respond to items on an 11-point rating scales, with ratings below 5 indicating that supervisory standards were not met. The ERS contains 35 items, although some items also allow for a response option of "Not Applicable". Items are added to produce a total score which is then divided by the number of items the respondent answered, producing an average score ranging between 0 and 10.

Tasks and Changes in Supervision Form (TCSF). The TCSF is also an original questionnaire designed by the first author and measures supervisory competence and changes in supervisory practice. It is a 53-item self-report questionnaire, in which supervisors indicate whether they believe they have sufficient or insufficient knowledge of skills. The questionnaire was administered before and after the training workshop. At post-training, supervisors were also asked whether any change had occurred in their practice since the workshop. Supervisors indicated change on five-point rating scales.

The TCSF was scored by calculating the proportion of items in which supervisors reported they had sufficient skills. Similar to the ERS, although the TCSF has been used regularly as part of a national supervisor training program, as yet there has been no investigation of the psychometric properties of this measure.

Evaluator Quality Rating. A single-item was used to estimate the overall competency of the supervision session. The evaluator provided a rating, scored out of 100, for each supervision session. Descriptions of score bands were agreed upon between the evaluator and researchers, with scores less than 50 indicating inadequate supervision competency.

Supervisory Confidence. A single, original item was used to measure supervisory confidence ("How confident do you feel about your effectiveness as a supervisor?" on a 10-point rating scale).

Supervisor Training Intervention

The first author, a clinical psychologist who has facilitated over 50 supervision-based workshops and trained over 1500 supervisors with consistently high evaluations, ran the supervision training program. The program consisted of a 2-day group workshop and feedback on three recorded supervision sessions. The workshop was based on literature regarding essential aspects to include in supervision training workshops (e.g. Falender, Shafranske, & Ofek, 2014; Watkins, 2012).

Procedure

Ethical approval was obtained from the associated university human research ethics committee. An invitation email was sent to all supervisors within the organisation. Participants were informed that regardless of participation in the research, all supervisors would have access to the training workshop. Participants were informed that their workplace would not be notified if their video supervision sessions did not pass the evaluation or if they chose not to participate in the training.

At each assessment time point supervisors were asked to record a supervision session with a consenting supervisee. Both supervisors and supervisees then completed self-report questionnaires at the end of the supervision session. The questionnaires and session recording were mailed to researchers, with the recording then sent to an evaluator for review. The evaluator was a psychologist with 8 years of experience evaluating supervisor effectiveness. The evaluator was trained by the first author to assess sessions using the ERS for the purposes of an ongoing supervision training program. The first author second-marked a random sample of 8 sessions to collaborate the marking by the evaluator. Evaluator feedback on time 1 and 2 tapes were provided after time 2 (just before the training workshop) included comments on supervision competencies as described in the ERS. Feedback on the third video-tape (after training) included comments on supervision competencies as described in

the ERS and the TCSF. The evaluator and second marker were not blind to the timing of recordings in relation to the training workshop.

Results

Twelve supervisors provided data at the first two time points and 10 at time point 3. Supervisors provided recordings and questionnaires using different supervisees at each time point. As such, analysis of supervisee data across all three time points was not possible. Rather time 2 and 3 data were used for analysis.

Given the brevity of this report, null results are not reported in detail, though it should be noted that null results with such a small sample may reflect Type II Error. Since the small sample size also affects the significant results *post hoc* power estimates are included for each significant result. Power analyses were typically above .8 (denoted 1-β). The reader can therefore evaluate the reliability of the significant results presented given the emergent effect sizes. Descriptive statistics for all measures are displayed in Table 1. No significant change over time was identified on the TCSF, SWAI, LASS, SSQ, or Supervisor Confidence.

INSERT TABLE 1 HERE

ERS

No effects were evident on the Supervisor Reports on the ERS. However, there was an effect on the Evaluator ERS Report. Across the baseline period mean scores on supervisory competence were around 7 out of 10. A significant effect of time was found ($F(1.129, 10.157) = 6.120$, $p = .030$, $f = .405$, 1-β=.82), which was followed up with within subjects t-tests using a Bonferroni correction ($\alpha_{adj} = .017$). One significant difference was found. Evaluator ERS ratings of supervisor competence significantly increased from pre to post intervention (time points 2 to 3; $t(9) = -3.481$, $p = .007$, $d = -.946$, 1-β-.99), whilst

remaining stable across the baseline (time 1 and 2) period. The significant difference in evaluator scores between time 2 and 3 was equivalent to an improvement (change index) of approximate 32.79%, and resulted in supervisors being rated as on average 8 out of 10 on the scale. The evaluator rated four supervisors at time point 1, five supervisors at time point 2, and one supervisor at time 3, as performing below competence on at least one item on the ERS.

Evaluator Quality Rating

A significant effect of time was found on the Evaluator Quality Rating ($F(2, 18) = 4.317$, $p = .029$, $f = .324$, 1-β=.63), which was followed up with within subjects t-tests using a Bonferroni correction ($\alpha_{adj} = .017$). At the adjusted alpha level, no pairwise comparisons were found to be significant. Using a 50% cut off score for overall competency, three supervisors scored below competency at time point 1, four scored below competency at time point 2, and one supervisor scored below competency at time point 3.

SRS

Mean subscale SRS supervisor self-report scores were above 7 out of 10, but the Total SRS scores at times 2 and 3 were below 36 out of 40, suggesting that the supervisors thought alliance was low at these time points. A one-way ANOVA revealed a significant difference in supervisor SRS total score across time (F (2, 18) = 6.342, $p = .008$, $f = .390$, 1-β=.79). This effect was followed up with post hoc within subjects t-tests, utilising a Bonferroni correction. Supervisor SRS total scores were significantly lower at time 2 ($t(9) = 3.584$, $p = .006$, $d = 1.135$, 1-β=.99) and 3 ($t(9) = 3.498$, $p = .001$, $d = 1.135$, 1-β=.99) than at time 1 (change index = 37.18% decrease in scores from time 1 to 2 and 3). By contrast, neither the Evaluator SRS total score, nor the Supervisee SRS total score showed significant time effects.

Satisfaction ratings of training program

The mean supervisor satisfaction rating with the training workshop was high (for both content (4.5 out of 5) and the Trainer (4.8 out of 5).

Discussion

The current study tested the effectiveness of a supervision training program on supervisor, supervisee, and evaluator perceptions of supervisory competence, confidence, satisfaction, and alliance. As hypothesized, supervisor competencies generally remained stable across the two baseline measurements and increased after the supervision training program (between time 2 and 3) according to the evaluator reports. Self-reported alliance decreased across time on one of the alliance measures (SRS), but was not noted by supervisees or evaluator on that same measure, and was not noted on the other two measures of alliance (LASS, SWAI). Lastly, supervisors positively evaluated the training program and stated that they had made positive changes in their supervision practise as a result of training.

This study collected ratings of supervisor competency from three different informants, and included both self-report and objective measures of competency. While some outcomes were similar across informants, some important differences occurred. Specifically, evaluator reports of mean supervision competency (ERS), and the ‘overall’ subscale of the alliance (SRS) was rated lower by the evaluator relative to supervisor and supervisee. These differences highlight the value of having multiple informants and using objective measures of competency and alliance. Without the evaluator reports there would be a perception that supervisor competencies and alliance were already high at baseline and training costs unlikely to be justifiable. However, by having access to the evaluator scores it is evident that even experienced supervisors can benefit from supervision training, and be objectively evaluated to have improved supervisor competencies (on the ERS) and overall quality of supervision (EQR).

While supervisors did not report that supervision training improved their competencies on the ERS, they did report an increase in skill use after intervention on the TCSF. Supervisors reported improvements in goal-setting, agenda setting, review, and seeking supervisee feedback in attempts to ensure supervisee needs were being met. These improvements are in line with Milne et al. (2011)'s conclusion that supervision training develops supervisor competencies in skill acquisition and behavioural change (e.g. ability to adjust behaviour to the needs of the supervisee).

The interpretation of the supervisor's self-reported decline in supervisory alliance across time is difficult given that it was not replicated on the other alliance measures, and not reported by supervisees or the evaluator. It may be that supervisors became more self-conscious and harsher critics of the alliance due to participation in the study.

Limitations

Four notable limitations of the current study are: a) the supervisor competency measures (ERS and TCSF) have not yet been psychometrically investigated); b) the small sample size; c) the variation in supervisor-supervisee pairs with some supervisors providing taped sessions with the same supervisee and others with different supervisees; and d) the evaluator was not blind to the timing of video-taped assessments.

Conclusion

With competency-based training of supervisors now compulsory in many countries this pilot study shows that supervisor competencies are generally stable across short-periods of time (e.g. baseline assessments) and that supervision training increases evaluator scores on supervisor competencies, even among experienced supervisors providing regular supervision in routine practise.

Key Points

What is Already Known About this Topic

1. Supervisor training is now compulsory in many countries, however there is limited evidence regarding best practice for these training programs.
2. It has been demonstrated that some aspects are effective in supervision training, including using modelling/observational learning, educational role-plays and corrective feedback.
3. There is evidence that supervision makes a difference to supervisees' practice, and that client outcome is enhanced.

What this Topic Adds

1. This pilot investigation provides preliminary support for the value of supervisor training programs for psychology/counselling professionals.
2. Supervisor competencies improved after supervision training as reported by an external evaluator but not on self-report of supervisor competency.
3. Supervisors enjoyed the training and reported an improvement in some areas of supervision skills.

References

Bambling, M., King, R., Raue, P., Schweitzer, R., & Lambert, W. (2006). Clinical supervision: Its influence on client-rated working alliance and client symptom reduction in the brief treatment of major depression. *Psychotherapy Research, 16*(3), 317-331. doi: 10.1080/10503300500268524.

Bargmann, S. & Robinson, B. (2012). Manual 2: Feedback-Informed clinical Work: The Basics. International Center for Clinical Excellence

Duncan, B. L., Miller, S., Sparks, J. A., Claud, D. A., Reynolds, L. R., Brown, J., & Johnson, L. D. (2003). The Session Rating Scale: Preliminary psychometric properties of a "working" alliance measure. *Journal of Brief Therapy, 3*(1), 3-12.

Efstation, J. F., Patton, M. J., & Kardash, C. M. (1990). Measuring the working alliance in counselor supervision. *Journal of Counseling Psychology, 37*(3), 322. doi: 10.1037/0022-0167.37.3.322

Falender, C. Shafranske, E. & Ofekb, A. (2014). Competent clinical supervision: Emerging effective practices. *Counselling Psychology Quarterly, 27(4), 393-408.* doi: 10.1080/09515070.2014.934785

Kavanagh, D. J., Spence, S., Sturk, H., Strong, J., Wilson, J., Worrall, L., …Skerrett, R. (2008). Outcomes of training in supervision: Randomised controlled trial. *Australian Psychologist, 43*(2), 96-104. doi: 10.1080/00050060802056534

Ladany, N., Hill, C. E., Corbett, M. M., & Nutt, E. A. (1996). Nature, extent, and importance of what psychotherapy trainees do not disclose to their supervisors. *Journal of Counseling Psychology, 43*(1), 10-24. doi: 10.1037/0022-0167.43.1.10

McMahon, M., & Simons, R. (2004). Supervision training for professional counselors: An exploratory study. *Counselor Education and Supervision, 43*, 301-309.

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. doi: 10.1348/014466502163796

Milne, D. L., Sheikh, A. I., Pattison, S., & Wilkinson, A. (2011). Evidence-based training for clinical supervisors: A systematic review of 11 controlled studies. *The Clinical Supervisor, 30*, 53-71. doi: 10.1080/07325223.2011.564955

Milne, D. & Westerman, C. (2001). Evidence-based clinical supervision: Rationale and illustration. *Clinical Psychology & Psychotherapy, 8(6),* 444–457. doi: 10.1002/cpp.297

Milne, D., & Reiser, R. P. (2014).SAGE: A Scale for Rating Competence in CBT Supervision, in *The Wiley International Handbook of Clinical Supervision* (eds C. E. Watkins and D. L. Milne) pp. 402-4150. Oxford: John Wiley & Sons, Lt.

Spence, S. H., Wilson, J., Kavanagh, D., Strong, J., & Worrall, L. (2001). Clinical supervision in four mental health professions: A review of the evidence. *Behaviour Change, 18*(3), 135-155. doi: 10.1375/bech.18.3.135

Tryon, G. S., Blackwell, S. C., & Hammel, E. F. (2007). A meta-analytic examination of client-therapist perspectives of the working alliance. Psychotherapy Research, 17, 629–642. doi: 10.1080/10503300701320611

Wainwright, N. A. (2010). *The development of the Leeds Alliance in Supervision Scale (LASS): A brief sessional measure of the supervisory alliance.* The University of Leeds.

Watkins, C. E. (2012). Development of the psychotherapy supervisor: Review of and reflections on 30 years of theory and research. *American Journal of Psychotherapy, 66(1), 45-83.*

Whitman, S.M., Ryan, B. & Rubenstein, D.F. (2001). Psychotherapy supervisor training:

Differences between psychiatry and other mental health disciplines. *Academic Psychiatry, 25,* 156-161. doi: 10.1176/appi.ap.25.3.156

Table 1
Descriptive Statistics for the ERS, TCSF, Evaluator Quality Rating, Supervisor Confidence Item, SWAI, LASS, SRS and SSQ

Measure	*Time 1*	*Time 2*	*Time 3*
ERS - Supervisor	8.23 (.71)	8.47 (.54)	8.02 (.88)
ERS - Evaluator	6.94 (1.23)	6.77 (1.01)	7.73 (1.02)
TCSF		.86 (.05)	.89 (.08)
Evaluator Quality Rating	61.40 (18.24)	63.20 (16.73)	75.10 (15.22)
SRS - Supervisor	37.30 (2.36)	35.20 (2.49)	35.10 (2.85)
SRS – Evaluator	32.00 (5.50)	27.70 (6.68)	31.10 (6.67)
SRS – Supervisee		36.71 (6.21)	38.14 (2.27)
SWAI- Supervisor			
Rapport		6.41 (.61)	5.96 (.41)
Client Focus		5.60 (.41)	5.72 (.30)
Identification		5.80 (.39)	5.64 (.60)
SWAI – Supervisee			
Rapport		6.35 (.67)	6.54 (.51)
Client Focus		6.04 (.67)	6.37 (.65)
Supervisor Confidence	7.85 (.63)	7.83 (.82)	7.75 (.92)
SSQ	29.70 (2.87)	29.39 (3.57)	28.90 (3.38)
LASS - Supervisor			
Approach	75.17 (11.42)	76.08 (10.59)	85.00 (11.65)
Relationship	82.80 (6.86)	84.60 (7.77)	81.60 (5.80)
Meeting Needs	78.10 (11.59)	78.90 (11.44)	78.70 (7.44)
LASS Evaluator			
Approach	47.40 (31.31)	45.00 (33.42)	71.00 (27.51)
Relationship	75.30 (12.10)	74.30 (20.55)	77.80 (12.68)
Meeting Needs	52.10 (23.68)	56.70 (26.82)	67.70 (23.76)
LASS - Supervisee			
Approach		83.86 (12.28)	90.14 (6.31)
Relationship		92.86 (3.58)	89.00 (9.11)
Meeting Needs		88.43 (11.97)	87.71 (9.27)

Note. Standard deviations shown in parentheses. For supervisor and evaluator reports n = 10 at time points 1, 2, and 3. For supervisee reports n = 7 at time points 2 and 3.