Behavioural and Cognitive Psychotherapy, 2020, **48**, 383–394
doi:10.1017/S1352465819000687

BABCP
www.babcp.com

MAIN

Providing objective feedback in supervision in motivational interviewing: results from a randomized controlled trial

Maria Beckman[1,*], Lars Forsberg[2], Helena Lindqvist[1] and Ata Ghaderi[3]

[1]Centre for Psychiatric Research, Karolinska Institutet, and Stockholm Health Care Services, Stockholm County Council, Sweden, [2]MIC Lab AB, Stockholm, Sweden and [3]Department of Clinical Neuroscience, Division of Psychology, Karolinska Institutet, Stockholm, Sweden
*Corresponding author. Email: maria.beckman@ki.se

(Received 19 March 2019; revised 27 August 2019; accepted 17 September 2019; first published online 05 November 2019)

Abstract

Background: The effects of the use of objective feedback in supervision on the supervisory relationship and skill acquisition is unknown.
Aims: The objective of this study was to evaluate the effects of two different types of objective feedback provided during supervision in motivational interviewing (MI) on: (a) the supervisory relationship, including potential feelings of discomfort/distress, provoked by the supervision sessions, and (b) the supervisees' skill acquisition.
Method: Data were obtained from a MI dissemination study conducted in five county councils across five county councils across Sweden. All 98 practitioners recorded sessions with standardized clients and were randomized to either systematic feedback based on only the behavioral component of a feedback protocol, or systematic feedback based on the entire protocol.
Results: The two different ways to provide objective feedback did not negatively affect the supervisory relationship, or provoke discomfort/distress among the supervisees, and the group that received the behavioural component of the feedback protocol performed better on only two of the seven skill measures.
Conclusions: Objective feedback does not seem to negatively affect either the supervisor–supervisee working alliance or the supervisees' supervision experience. The observed differences in MI skill acquisition were small, and constructive replications are needed to ascertain the mode and complexity of feedback that optimizes practitioners' learning, while minimizing the sense of discomfort and distress.

Keywords: clinical supervision; motivational interviewing; objective feedback; training

Introduction

Efficient supervision as part of training is an important factor for dissemination and implementation of evidence-based treatments (EBT) into clinical practice (Edmunds *et al.*, 2013; Fixsen *et al.*, 2005). Motivational interviewing (MI) has been demonstrated as a clinically effective EBT and has been implemented in a number of different clinical and healthcare settings (Miller and Moyers, 2017). The client-centred and directional therapeutic method for strengthening clients' motivation and commitment to change requires careful training to achieve sustained practice change (Miller and Rollnick, 2013). However, prior research shows that training formats that include supervision are more likely to enhance skills than the most common form of MI training (i.e. a 2-day workshop) (Barwick *et al.*, 2012; de Roten *et al.*, 2013; Madson *et al.*, 2009; Schwalbe *et al.*, 2014).

© British Association for Behavioural and Cognitive Psychotherapies 2019

https://doi.org/10.1017/S1352465819000687 Published online by Cambridge University Press

Clinical supervision

Although clinical supervision is widely recognized as important for professional development and to ensure client outcomes, less is known about the impact of supervision on both clinical practice and client outcome (Bambling *et al.*, 2006). However, research has shown that supervision can be beneficial for the supervisee's self-awareness, skills, self-efficacy, theoretical orientation, and for the supervisee–client relationship (Watkins, 2011; Wheeler and Richards, 2007). There is also growing evidence for the association between supervision and better outcomes for both clinicians and their clients (Rakovshik *et al.*, 2016). However, the sparse previous studies on the effect on therapist behaviour and client outcome shows ambiguous results (Simpson-Southward *et al.*, 2016; Zarbock *et al.*, 2009), and the methodological shortcomings of this research make it difficult to identify the specific supervision features that constitute best practices (Bearman *et al.*, 2017). In recent years, there has been an increased focus on ensuring supervisor competence and effective supervision practice in both evidence-based clinical supervision (Milne and Reiser, 2012) and competency-based supervision (Falender and Shafranske, 2007). Both these supervision practices share a few core components: articulations of supervision goals, feedback on both adherence and skills based on monitoring of sessions, individualized coaching/training including active learning (e.g. structured performance feedback, reflection, and behavioural rehearsal/role-play), and evaluation of targeted supervision competencies. Despite this increased attention to supervision as part of training programmes for practitioners, many questions remain regarding the included strategies for a successful transfer into clinical settings (Edmunds *et al.*, 2013).

Systematic feedback

Performance-related feedback is perceived as an important part of clinical supervision and has demonstrated slight but possibly essential improvements in clinical practice (Ivers *et al.*, 2012). Feedback typically contains evaluation of a supervisee's skills and adherence (Parsons *et al.*, 2012). Feedback based on monitoring of sessions often also informs the supervisees how well they performed relative to a standard level, allows for the supervisor to direct the supervisee's attention to specific behaviours, and may also aid the supervisee to self-assess more accurately (Parsons *et al.*, 2012). The Motivational Interviewing Treatment Integrity (MITI) Code is a coding system with acceptable psychometric properties (Moyers *et al.*, 2005), intended to be used as a treatment integrity measure in clinical trials of MI, and for providing structured feedback to practitioners in non-research settings. MITI 3.1 (Moyers *et al.*, 2010) has two components: (1) the behaviour counts (i.e. giving information, MI adherent behaviours, MI non-adherent behaviours, closed questions, open questions, simple reflections and complex reflections), and (2) the global scores, which capture the overall impression of the session in five different dimensions (i.e. empathy, evocation, collaboration, autonomy and direction) on a Likert-type scale from 1 to 5. MITI also specifies proficiency and competency thresholds for practitioners (Table 1). When used as a feedback tool, the MITI provides a comprehensive picture of MI skills to the supervisee, with the primary purpose of supporting the acquisition of skills. However, in a recent semi-structured interview with ten MI supervisors using the MITI for objective feedback (Beckman *et al.*, 2017a), 50% of the supervisors reported that the MITI feedback, and in particular the global dimensions, may have induced negative emotions among the supervisees that hindered their learning. The supervisors perceived the behaviour frequency counts as being more objective and less judgemental, and argued that especially supervisees with longer work experience had difficulties handling the global scores on the lower side. This viewpoint is consistent with earlier research which suggests that supervisees may fear negative feedback (Abernethy and Cook, 2011; Bernard and Goodyear, 2014; Clarke and Giordano, 2013; Ellis *et al.*, 2015; Friedberg *et al.*, 2009; Lombardo *et al.*, 2009), and that supervisors often feel critical and

https://doi.org/10.1017/S1352465819000687 Published online by Cambridge University Press

Table 1. Recommended MITI beginning proficiency and competency thresholds

MITI summary scores	Beginning proficiency	Competency
Global clinician ratings[1]	Average of 3.5	Average of 4
Reflection to question ratio (R:Q)[2]	1	2
Per cent open questions (%OC)[3]	50%	70%
Per cent complex reflections (%CR)[4]	40%	50%
Per cent MI-adherent (% MIA)[5]	90%	100%

MITI, Motivational Interviewing Treatment Integrity Code. [1](Empathy + Evocation + Collaboration + Autonomy)/4. [2](Simple + Complex reflections)/(Open + Closed questions). [3]Open questions/(Open + Closed questions). [4]Complex reflections/(Simple + Complex reflections). [5]MI Adherent behaviours/(MI adherent + MI non-adherent behaviours).

worry that their feedback may harm the supervisory working alliance (Chur-Hansen and McLean, 2006). Self-reported data even suggests that supervisors sometimes withhold corrective feedback and/or give higher ratings to avoid negative reactions or harming the supervisory relationship (Turner *et al.*, 2016). High levels of supervisee anxiety have been proposed to cause defensiveness, reduce readiness to reveal information, cause supervisor–supervisee role conflicts, and decrease supervisee clinical performance (Ellis *et al.*, 2015). Findings from the science of learning have also suggested that aspects of emotions could be important to consider (La Rochelle *et al.*, 2011), and emotions have thus gained increased attention in recent years' medical education research (Young *et al.*, 2014). Emotions affect both the attention to information and how memory is stored and retrieved (Gooding *et al.*, 2017): positive emotions during learning are linked to deeper cognitive processing and improved learning, while negative emotions are linked to more superficial cognitive processing and impeded learning (Young *et al.*, 2014). However, other findings from the supervision field propose that the usual assumption of recordings of sessions as overburdening is a misconception, and that supervisees often quickly adapt to recordings of sessions and can handle corrective feedback, especially within a positive supervisory working alliance (Ellis, 2010; Ladany *et al.*, 2013). Some researchers even suggest that providing both positive and corrective feedback enhances the supervisory relationship and increases the supervisor's satisfaction with their role (Chur-Hansen and McLean, 2006).

The aim of this study was to evaluate how two different types of objective feedback (i.e. based on only the behaviour counts part of the MITI, or based on both the five global dimensions and the behaviour counts of the MITI) provided during six sessions of MI supervision affect the supervisory relationship, potential feelings of discomfort/distress provoked by the supervision sessions, and the supervisees' MI skill acquisition.

Method

Data were obtained from a Swedish MI dissemination study conducted in five county councils across Sweden (ClinicalTrials.gov: NCT01197027). In the original study, a total of 174 self-selected practitioners were randomized to either regular county council MI training or regular county council MI training with six sessions of subsequent supervision. All the MI training increased the participants' MI skills to about the same level, but the group that received subsequent supervision had gained more MI skills at follow-up (Beckman *et al.*, 2017b). Analyses also showed generally preserved MI proficiency levels for all participants at follow-up, but the majority did not reach the beginning proficiency levels at either the post-workshop or at the follow-up assessment (Beckman *et al.*, 2017b).

https://doi.org/10.1017/S1352465819000687 Published online by Cambridge University Press

Table 2. Demographic characteristics of the two groups

Group	MITI BC	MITI GD + BC
Age, mean (*SD*)	46.3 (11.3)	42.6 (10.7)
Female, *n* (%)	35 (89.7)	36 (81.8)
Bachelor's degree, *n* (%)	24 (61.5)	27 (61.4)
Master's degree, *n* (%)	13 (33.4)	12 (27.2)
Doctorate degree, *n* (%)	2 (5.1)	5 (11.4)
Professions		
Nurse	12	11
Social worker	7	6
Physiotherapist	3	8
Physician	3	6
Dietician	2	3
Educationalist	3	2
Counsellor	3	1
Psychologist	3	4
Other	3	3
Missing	8	7

MITI BC, supervision including systematic feedback based on only the behaviour counts part of the MITI; MITI GD + BC, supervision including systematic feedback based on both the five global dimensions and the behaviour counts; *SD*, standard deviation.

Participants

The sample consisted of the 98 participants from the dissemination study randomized to receive MI supervision. All participants participated in MI workshop trainings in five Swedish county councils from January 2013 to September 2014. The mean age was 44.3 years ($SD = 11.1$), and 85.5% were female ($n = 71$). The education level varied from bachelor's degree ($n = 51$, 61.4%) to master's degree ($n = 25$, 30.2%) and doctorate degree ($n = 7$, 8.4%), and the participants came from a variety of professions (Table 2).

Procedure

Written informed consent was obtained from all participants before they enrolled (the Regional Ethical Review Board in Stockholm, Sweden; 2012/2195-31/5). The participants were then randomized to one of two groups (see Fig. 1): (1) supervision including systematic feedback based on only the behaviour counts part of the MITI (MITI BC), or (2) supervision including systematic feedback based on both the five global dimensions and the behaviour counts of the MITI (MITI GD + BC). The randomization procedure was conducted across all participants, and are described in detail elsewhere (Beckman *et al.*, 2017b). The participants recorded seven 20-minute telephone sessions with actors playing standardized clients: the first one after the MI training (baseline), and the following with monthly intervals. The supervision sessions followed after sessions 1 to 6. Before all recordings, the participants were sent information regarding the role-played client and a specified session target behaviour by email. The standardized clients were based on the lifestyle habits in The Swedish National Guidelines for Methods of Preventing Disease (socialstyrelsen.se/nationalguidelines): hazardous use of alcohol, insufficient physical activity, unhealthy eating habits, and tobacco use.

Telephone supervision

The study's six monthly sessions of individual telephone supervision were all 30 minutes long and performed by twelve coders trained in the manual based MI supervision at the Motivational

https://doi.org/10.1017/S1352465819000687 Published online by Cambridge University Press

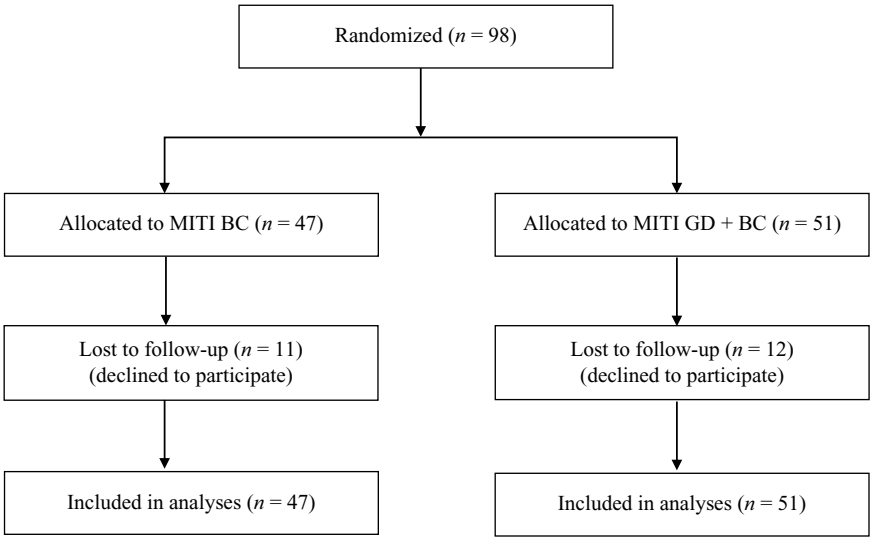


Figure 1. Flow diagram. MITI BC = supervision including systematic feedback based on only the behavior counts part of the MITI; MITI GD + BC = supervision including systematic feedback based on both the five global dimensions and the behaviour counts of the MITI.

Interviewing Quality Assurance (MIQA) group at Karolinska institutet in Sweden. The primary purpose of the sessions was further development of MI skills. The MITI BC group received their systematic feedback based on only the behaviour counts part of the MITI. In these sessions, the supervisor could speak in general terms about the global dimensions (e.g. state that the supervisee had a relatively large proportion of complex reflections which could affect the global variable empathy). However, the global scores had been removed from the protocol, and the supervisor did not talk about the supervisee's individual global scores in these sessions. The MITI GD + BC group received feedback based on the results of both the five global dimensions and the behaviour counts. Ten minutes before the sessions, the coder emailed the protocol to the participant. All sessions were manual-based (please see 'Supplementary material'), and are described in detail elsewhere (Beckman *et al.*, 2017b).

Assessment

Participant's demographics (Table 2) were collected with a self-report questionnaire. The Swedish version of the MITI, version 3.1.1 (Forsberg *et al.*, 2011) was used for assessing the recordings for MI proficiency. The twelve coders had all received 120 hours of initial coding training and participated in weekly group-codings to maintain high inter-rater reliability. The coders were not blinded for group allocation. To measure the supervisory relationship, an adapted version of the short Working Alliance Inventory (WAI-S), and a question regarding discomfort/distress were used. The WAI (Horvath and Greenberg, 1989) is one of the questionnaires most frequently used to measure working alliance (Busseri and Tyler, 2003). The WAI-S (Tracey and Kokotovic, 1989) contains 12 items rated on a 7-point scale. The total score measures the global working alliance, and three subscales assess primary components of the working alliance: Goal (agreement with regard to goals), Task (agreement with regard to tasks) and Bond (the empathic bond between client and therapist). The adaptation consisted of minor reframing of items to apply to the relationship between supervisee–supervisor instead of client–therapist. Potential feelings of discomfort/distress provoked by the supervision sessions were assessed with one single question – the participants were asked to

https://doi.org/10.1017/S1352465819000687 Published online by Cambridge University Press

specify a number from 1 (i.e. *No feeling of discomfort/distress or anything that may fall under the category of a negative feeling or experience*), to 10 (i.e. *A feeling of discomfort/distress or a negative feeling or experience of any kind*) that corresponded with their experience of the supervision session.

Data analysis

All data were analysed with the Statistical Package for the Social Sciences (SPSS), version 23.0. The primary outcomes were the seven MITI proficiency measurements, the WAI-S scores, and the single question regarding discomfort/distress at baseline and the 6-month assessment. To examine baseline differences, *t*-tests and chi-square tests were performed. To control for incomplete and nested data and to handle non-normal data efficiently, a generalized linear mixed model (GLMM) was used. Both main effects and interactions were examined for all primary outcomes. The best distribution for data was found by QQ-plots and other descriptive statistics. For the normal covariance structure, the identity link was used. For the gamma distribution, the log link was used. Adjustments other than nesting for repeated measures and random intercept for individuals did not provide a better data fit. Chi-square analyses were used for the proportions of clinicians that met the MITI beginning proficiency levels, the Bonferroni correction was used for reducing the chances of false-positive results, and Cohen's *d* was used as effect size measure. Moreover, to assess the MIQA coders' inter-rater agreement, 12 randomly selected recordings are two times a year double-coded by the whole coding group. Intraclass correlation coefficients (ICC) in a two-way mixed model with single measures and absolute agreement are then used for calculating the coders inter-rater agreement. According to Cicchetti's (1994) system, an ICC below 0.40 is considered poor, an ICC between 0.40 and 0.59 is considered fair, an ICC between 0.60 and 0.74 is considered good, and an ICC between 0.75 and 1.00 is considered excellent. In January 2014 (i.e. the middle of the study period), the MIQA ICCs for all the MITI variables ranged between good (Empathy 0.60, Evocation 0.69, Collaboration 0.74, Closed questions 0.64, Simple reflections 0.73, and Complex reflections 0.68), to excellent (Autonomy 0.75, Giving information 0.89, MI adherent behaviors 0.81, and Open questions 0.92), except for Direction and MI non-adherent, for which the results were fair (Direction 0.49, and MI non-adherent behaviours 0.59).

Results

Working alliance and the discomfort/distress question

The group means of WAI-S and the discomfort/distress question for the two groups at the two assessment points are presented in Table 3. The GLMM analyses of the working alliance showed no significant interaction or group effects, but a significant effect of time for the total WAI-S score as well as for two of the three subscales (Table 3). The analyses also showed a significant time and group interaction for the discomfort/distress question (Table 3). The MITI GD + BC supervision group had a somewhat higher score at baseline and a somewhat lower score at the 6-month assessment compared with the MITI BC group, indicating a more apparent decrease of discomfort/distress over time for the MITI GD+BC group. Furthermore, the analysis of the discomfort/distress question showed no significant group effect but a significant time effect (Table 3).

MI skill acquisition

The group means of the seven MITI proficiency measures for the two groups at the two assessment points are presented in Table 3. The GLMM analyses showed significant time and

https://doi.org/10.1017/S1352465819000687 Published online by Cambridge University Press

https://doi.org/10.1017/S1352465819000687 Published online by Cambridge University Press

Table 3. Group means of the seven MITI proficiency measures, the WAI-S and the discomfort/distress question for the two groups at the two assessment points

	Baseline mean (*SE*)		6-month mean (*SE*)	
	MITI BC	MITI GD + BC	MITI BC	MITI GD + BC
Empathy (1–5)	3.50 (0.12)	3.46 (0.11)	3.88 (0.14)	3.67 (0.12)
Time: ***F* (1,152) = 8.52**, = **.004**, ***d* = 0.59**; Group: *F* (1,152) = 0.69, *p* = .408, *d* = 0.17; Time×Group: *F* (1,152) = 0.61, *p* = .436, *d* = 0.16				
MI Spirit (1–5)	3.20 (0.11)	3.15 (0.10)	3.54 (0.12)	3.71 (0.13)
Time: ***F* (1,152) = 15.99**, ***p* = .001**, ***d* = 0.8**; Group: *F* (1,152) = 0.27, *p* = .603, *d* = 0.10; Time×Group: *F* (1,152) = 1.09, *p* = .297, *d* = 0.21				
Adherent behaviours	2.12 (0.23)	2.19 (0.20)	3.97 (0.36)	3.85 (0.32)
Time: ***F* (1,120) = 40.51**, ***p* = .001**, ***d* = 1.29**; Group: *F* (1,120) = 0.00, *p* = .983, *d* = 0.00; Time×Group: *F* (1,120) = 0.12, *p* = .735, *d* = 0.07				
Non-adherent behaviours	0.89 (0.21)	0.91 (0.18)	0.22 (0.10)	1.15 (0.23)
Time: ***F* (1,152) = 6.33**, ***p* = .013**, ***d* = 0.51**; Group: ***F* (1,152) = 5.19**, ***p* = .024**, ***d* = 0.46**; Time×Group: ***F* (1,152) = 12.34**, ***p* = .001**, ***d* = 0.71**				
Reflection to question	1.02 (0.10)	0.89 (0.08)	1.02 (0.10)	0.90 (0.09)
Time: *F* (1,151) = 0.00, *p* = .958, *d* = 0.00; Group: *F* (1,151) = 1.27, *p* = .262, *d* = 0.23; Time×Group: *F* (1,151) = 0.00, *p* = .947, *d* = 0.00				
% Open questions	46.10 (2.98)	40.41 (2.77)	43.21 (3.27)	48.15 (2.93)
Time: *F* (1,152) = 0.75, *p* = .387, *d* = 0.17; Group: *F* (1,152) = 0.02, *p* = .900, *d* = 0.03; Time×Group: *F* (1,152) = 3.78, *p* = .054, *d* = 0.39				
% Complex reflections	35.54 (2.75)	43.92 (3.00)	52.33 (3.66)	50.22 (3.43)
Time: ***F* (1,148) = 20.33**, ***p* = .001**, ***d* = 0.91**; Group: *F* (1,148) = 0.72, *p* = .396, *d* = 0.17; Time×Group: ***F* (1,148) = 4.29**, ***p* = .040**, ***d* = 0.42**				
WAI-S Total (1–7)	5.63 (0.70)	5.61 (0.83)	5.96 (0.78)	6.19 (0.59)
Time: ***F* (1,137) = 19.32**, ***p* = .000**, ***d* = 0.89**; Group: *F* (1,137) = 0.41, *p* = .526, *d* = 0.13; Time×Group: *F* (1,137) = 1.32, *p* = .252, *d* = 0.23				
WAI-S Goal (1–7)	5.99 (0.10)	5.92 (0.10)	6.10 (0.11)	6.20 (0.10)
Time: *F* (1,142) = 4.70, *p* = .052, *d* = 0.44; Group: *F* (1,142) = 0.02, *p* = .880, *d* = 0.03; Time×Group: *F* (1,142) = 0.77, *p* = .383, *d* = 0.18				
WAI-S Task (1–7)	5.50 (0.14)	5.61 (0.13)	5.90 (0.15)	6.16 (0.14)
Time: ***F* (1,144) = 17.90**, ***p* = .000**, ***d* = 0.98**; Group: *F* (1,144) = 1.35, *p* = .246, *d* = 0.23; Time×Group: *F* (1,144) = 0.47, *p* = .496, *d* = 0.14				
WAI-S Bond (1–7)	5.36 (0.16)	5.25 (0.14)	5.81 (0.18)	6.09 (0.18)
Time: ***F* (1,139) = 24.00**, ***p* = .000**, ***d* = 1.13**; Group: *F* (1,139) = 0.15, *p* = .696, *d* = 0.08; Time×Group: *F* (1,139) = 1.21, *p* = .139, *d* = 0.22				
Discomfort/distress (1–10)	2.13 (0.23)	2.64 (0.22)	2.01 (0.25)	1.59 (0.24)
Time: ***F* (1,142) = 7.21**, ***p* = .008**, ***d* = 0.62**; Group: *F* (1,142) = 0.03, *p* = .862, *d* = 0.03; Time×Group: ***F* (1,142) = 4.57**, ***p* = .034**, ***d* = 0.49**				

MITI, Motivational Interviewing Treatment Integrity Code; WAIS-S, Working Alliance Inventory (Short Form); *SE*, standard error; MITI BC, supervision including systematic feedback based on the behaviour counts part of the MITI; MITI GD + BC, supervision including systematic feedback based on both the five global dimensions and the behaviour counts.

Table 4. The number and percentage of the participants in the two groups reaching the MITI beginning proficiency thresholds at the two assessment points, and statistical comparisons between the groups

	n (%) Proficient at baseline	φ	*n* (%) Proficient at 6 months	φ
Global clinician ratings				
MITI BC	18 (50.0)	.01	15 (41.7)	.01
MITI GD + BC	19 (48.7)		16 (41.0)	
Reflection to question ratio				
MITI BC	15 (41.7)	.03	21 (58.3)	.23
MITI GD + BC	15 (38.5)		14 (38.5)	
Per cent open questions				
MITI BC	17 (47.2)	.12	16 (44.4)	.04
MITI GD + BC	14 (35.9)		19 (48.7)	
Per cent complex reflections				
MITI BC	18 (50.0)	.06	28 (77.8)	.04
MITI GD + BC	22 (56.4)		29 (74.4)	
Per cent MI-adherent				
MITI BC	23 (63.9)	.02	28 (77.8)	.28
MITI GD + BC	24 (61.5)		20 (51.3)	

φ, phi coefficient; MITI BC, supervision including systematic feedback based on only the behavior counts part of the MITI; MITI GD + BC, supervision including systematic feedback based on both the five global dimensions and the behaviour counts.

group interactions for MI non-adherent behaviours and per cent complex reflections (Table 3). The MITI BC supervision group performed better at the 6-month assessment with regard to both of those proficiency measures, and thus had a better development over time. Moreover, the analyses also showed one significant group effect for MI non-adherent behaviours, and significant time effects for five of the seven MITI proficiency measures (Table 3).

Table 4 shows the number and proportion of participants in the two groups reaching MITI beginning proficiency thresholds at the two assessment points. At baseline, directly after the county council workshop trainings, 35.9– 63.9% of the participants reached beginning proficiency levels on indicators. At the 6-month assessment, the proportion had increased for both groups for per cent complex reflections, and decreased for global clinician ratings. For the other indicators, the results for the two groups were mixed at the 6-month assessment. The MITI BC group had a higher proportion of participants reaching beginning proficiency levels on four of the five indicators at the 6-month assessment, but after the Bonferroni correction was applied none of these differences remained significant (Table 4). Per cent MI-adherent showed the highest proportion of participants reaching beginning proficiency thresholds baseline, and per cent complex reflections showed the highest proportion at the 6-month assessment.

Discussion

The objective of study was to evaluate how two different types of objective MI feedback affect the supervisory relationship, potential feelings of discomfort/distress provoked by the feedback, and the supervisees' skill acquisition. Neither of the two different ways to provide objective feedback seemed to negatively affect the supervisory relationship: at the 6-month assessment, the WAI-S total score and all three subscales (i.e. goal, task and bond) had increased for both groups from already high scores at baseline (Table 3), with no between-groups differences. However, the results did reveal a significant difference in discomfort/distress between the two groups. Nonetheless, at the 6-month assessment, the discomfort/distress scores had decreased in both groups from already

https://doi.org/10.1017/S1352465819000687 Published online by Cambridge University Press

low scores at baseline (Table 3), indicating that the objective feedback was not difficult to handle for any of the supervisees. The differences between the MI skill acquisition of the two feedback groups were also small (i.e. the MITI BC group performed better on only two of the seven skill measures).

Working alliance and the discomfort/distress question

With slightly higher scores at the first recording (indicating a somewhat higher discomfort/distress) and slightly lower scores at the last recording (Table 3), the MITI GD+BC group had a more apparent decrease over time on the discomfort/distress question. There is no obvious explanation for these results. Maybe the exposure to the global dimensions of the MITI led to an extinction of the initial provoked discomfort/distress for the MITI GD+BC group over the course of time? Exposure techniques are common in effective psychotherapies and involve repeated exposure to a feared object or situation in order to overcome distress (Vervliet *et al.*, 2013). However, the difference between the groups was small and measured with just one single question. Additionally, the group effect for the discomfort/distress question was not significant and the low scores for both groups at both assessment points indicates that the objective feedback did not provoke significant negative emotions in any of the groups, at any of the time points. These results are consistent with a recent study that compared the effect of active learning techniques versus 'supervision as usual', where participants in both groups reported equally high levels of satisfaction (Bearman *et al.*, 2017), and with previous conclusions that supervisees can handle both positive and corrective performance feedback, especially within a positive supervisory relationship (Ellis, 2010; Ladany *et al.*, 2013).

MI skill acquisition

The differences between the groups' MI skill acquisition were small (i.e. the MITI BC group performed better on only two of the seven skill measures), and it is not clear what really generated them. In the study by Beckman and colleagues (2017), the supervisors expressed concern that in particular the global dimensions of the MITI may have caused negative emotions that hindered the supervisees' learning. As neither of the two different ways to provide objective feedback in this study caused negative emotions or negatively affected the supervisory relationship, the concern expressed by the supervisors in the previous study (Beckman *et al.*, 2017a) might be dismissed. The MITI 3.1 is a multi-faceted compilation of supervisees' skills (Moyers *et al.*, 2010), and the minor differences in skill acquisition between the two groups may simply be explained by the limited number of variables presented at each session for the MITI BC group. Cognitive load theory (CLT) (van Merrienboer and Sweller, 2010) is one of the leading cognitive learning theories with implications for both complex learning and skills acquisition. CLT provides a theoretical framework of the cognitive architecture, with the basic assumption that the capacity of the working memory is limited. According to the theory, three sources of cognitive load should be considered in all learning situations: the *intrinsic load* from the complexity of the learning task, the *extraneous load* created by the learning situation, and the *germane load* from processing the material. When the cognitive load exceeds the capacity of the working memory, learning and skills acquisition can be reduced or even hindered (van Merrienboer and Sweller, 2010). Limiting the number of variables when providing objective feedback could possibly reduce the risk of overloading supervisees' cognitive capacity and thus facilitate learning during supervision. Part-task, sequenced, spaced or distributed practice also allows for repetition and for the supervisee to reflect and practise in between sessions (Kerfoot *et al.*, 2007; Rakovshik and McManus, 2010; Young *et al.*, 2014). Acquisition of long-term knowledge and skills is dependent on all these variables, and training provided in doses could thereby possibly be a more effective form of learning.

https://doi.org/10.1017/S1352465819000687 Published online by Cambridge University Press

Limitations

This study has several limitations: The self-selected participants in the Swedish MI dissemination study may limit the generalizability of the results. Also, the design of the dissemination study did not allow a third supervision group (control) without objective feedback, and the recruitment difficulties in that trial had consequences for the ideal sample size. Moreover, a larger number of recordings at each assessment point could have provided more accurate estimates of MI skills, as MI performance can vary significantly within the same practitioner (Dunn *et al.*, 2016; Imel *et al.*, 2014). The use of standardized clients can also provide inadequate information regarding MI in actual clinical practice (Decker *et al.*, 2014). However, standardized clients keep the characteristics fixed among participants and assessment points, and if not used simultaneously with clients, role-plays can provide valuable data on MI performance (Decker *et al.*, 2013; Imel *et al.*, 2014). Furthermore, during coding, the coders were not blind to the group allocations, and the same coder both coded the recordings and provided supervision, which may have affected the reliability of the codings (Moyers *et al.*, 2016). Despite these limitations, this study can provide some knowledge and direction for future supervision studies, by being one of few studies that compare two different way to provide feedback based on practice samples.

Conclusions

Pursuant to earlier research (Ellis, 2010; Ladany *et al.*, 2013), objective feedback does not seem to negatively affect either the supervisor–supervisee working alliance or the supervisees' supervision experience. Although restricting the number of variables when providing objective feedback might promote learning during supervision, the observed differences between the two groups' MI skill acquisition were small and on only two of seven parameters. Earlier research has found providing feedback as one of the weaker skill areas for clinical supervisors (Cummings *et al.*, 2015), and lack of feedback or exclusively positive feedback is a frequent criticism during clinical training (Cummings *et al.*, 2015). Directions for future practice could be focused on both positive and corrective feedback, based on monitoring of sessions, as important parts of clinical supervision. Directions for future research could instead be focused on comparing different ways of providing feedback. As objective feedback seem to be an important part of supervision (Ivers *et al.*, 2012), and efficient supervision an important factor for the dissemination and implementation of EBT (Edmunds *et al.*, 2013; Fixsen *et al.*, 2005), constructive replications are needed to ascertain the mode and complexity of feedback that optimizes practitioners' learning, while minimizing the sense of discomfort and distress.

Supplementary material. To view supplementary material for this article, please visit: https://doi.org/10.1017/S1352465819000687

Acknowledgements. The authors would like to thank the research participants.

Financial support. The Swedish National Board of Health and Welfare funded this study (2.4-59457/2012).

Conflicts of interest. The authors have no conflicts of interest with respect to this publication.

Ethical statements. The authors have abided by the Ethical Principles of Psychologists and Code of Conduct as set out by the American Psychological Association. Ethical approval for the trials was granted by the Regional Ethical Review Board in Stockholm, Sweden; 2012/2195-31/5).

References

Abernethy, C., & Cook, K. (2011). Resistance or disconnection? A relational-cultural approach to supervisee anxiety and nondisclosure. *Journal of Creativity in Mental Health*, *6*, 2–14. doi: 10.1080/15401383.2011.560067

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

https://doi.org/10.1017/S1352465819000687 Published online by Cambridge University Press

Barwick, M. A., Bennett, L. M., Johnson, S. N., McGowan, J., & Moore, J. E. (2012). Training health and mental health professionals in motivational interviewing: a systematic review. *Children and Youth Services Review*, *34*, 1786–1795. doi: 10.1016/j.childyouth.2012.05.012

Bearman, S. K., Schneiderman, R. L., & Zoloth, E. (2017). Building an evidence base for effective supervision practices: an analogue experiment of supervision to increase EBT fidelity. *Administration and Policy in Mental Health*, *44*, 293–307. doi: 10.1007/s10488-016-0723-8

Beckman, M., Bohman, B., Forsberg, L., Rasmussen, F., & Ghaderi, A. (2017a). Supervision in motivational interviewing: an exploratory study. *Behavioural and Cognitive Psychotherapy*, *45*, 351–365. doi: 10.1017/S135246581700011X

Beckman, M., Forsberg, L., Lindqvist, H., Diez, M., Eno Persson, J., & Ghaderi, A. (2017b). The dissemination of motivational interviewing in Swedish county councils: results of a randomized controlled trial. *PloS One*, *12*, e0181715. doi: 10.1371/journal.pone.0181715

Bernard, J. M., & Goodyear, R. K. (2014). *Fundamentals of Clinical Supervision* (5th edn). Boston, MA, USA: Pearson.

Busseri, M. A., & Tyler, J. D. (2003). Interchangeability of the Working Alliance Inventory and Working Alliance Inventory, Short Form. *Psychological Assessment*, *15*, 193–197. doi: 10.1037/1040-3590.15.2.193

Chur-Hansen, A., & McLean, S. (2006). On being a supervisor: the importance of feedback and how to give it. *Australas Psychiatry*, *14*, 67–71. doi: 10.1080/j.1440-1665.2006.02248.x

Cicchetti, D. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*, 284–290.

Clarke, P. B., & Giordano, A. L. (2013). The motivational supervisor: motivational interviewing as a clinical supervision approach. *The Clinical Supervisor*, *32*, 244–259. doi: 10.1080/07325223.2013.851633

Cummings, J. A., Ballantyne, E. C., & Scallion, L. M. (2015). Essential processes for cognitive behavioral clinical supervision: agenda setting, problem-solving, and formative feedback. *Psychotherapy (Chic)*, *52*, 158–163. doi: 10.1037/a0038712

de Roten, Y., Zimmermann, G., Ortega, D., & Despland, J. N. (2013). Meta-analysis of the effects of MI training on clinicians' behavior. *Journal of Substance Abuse Treatment*, *45*, 155–162. doi: 10.1016/j.jsat.2013.02.006

Decker, S. E., Carroll, K. M., Nich, C., Canning-Ball, M., & Martino, S. (2013). Correspondence of motivational interviewing adherence and competence ratings in real and role-played client sessions. *Psychological Assessment*, *25*, 306–312. doi: 10.1037/a0030815

Decker, S. E., Nich, C., Carroll, K. M., & Martino, S. (2014). Development of the Therapist Empathy Scale. *Behavioural and Cognitive Psychotherapy*, *42*, 339–354. doi: 10.1017/S1352465813000039

Dunn, C., Darnell, D., Atkins, D. C., Hallgren, K. A., Imel, Z. E., Bumgardner, K., … & Roy-Byrne, P. (2016). Within-provider variability in motivational interviewing integrity for three years after MI training: does time heal? *Journal of Substance Abuse Treatment*, *65*, 74–82. doi: 10.1016/j.jsat.2016.02.008

Edmunds, J. M., Beidas, R. S., & Kendall, P. C. (2013). Dissemination and implementation of evidence-based practices: training and consultation as implementation strategies. *Clinical Psychology*, *20*, 152–165. doi: 10.1111/cpsp.12031

Ellis, M. V. (2010). Bridging the ccience and practice of clinical supervision: some discoveries, some misconceptions. *The Clinical Supervisor*, *29*, 95–116. doi: 10.1080/07325221003741910

Ellis, M. V., Hutman, H., & Chapin, J. (2015). Reducing supervisee anxiety: effects of a role induction intervention for clinical supervision. *Journal of Counseling Psychology*, *62*, 608–620. doi: 10.1037/cou0000099

Falender, C. A., & Shafranske, E. P. (2007). Competence in competency-based supervision practice: construct and application. *Professional Psychology: Research and Practice*, *38*, 232–240. doi: 10.1037/0735-7028.38.3.232

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). Implementation research: a synthesis of the literature. Available at: https://nirn.fpg.unc.edu/sites/nirn.fpg.unc.edu/files/resources/NIRN-MonographFull-01-2005.pdf

Forsberg, L., Forsberg, L., Forsberg, K., van Loo, T., & Rönnqvist, S. (2011). Motivational interviewing treatment integrity 3.1 (MITI 3.1): MITI kodningsmanual 3.1. Retrieved from: https://www.miclab.org/sites/default/files/files/MITI%203_1%20SV.pdf

Friedberg, R., Gorman, A., & Beidel, D. (2009). Training psychologists for cognitive-behavioral therapy in the raw world – a rubric for supervisors. *Behavior Modification*, *33*, 104–123. doi: 10.1177/0145445508322609

Gooding, H. C., Mann, K., & Armstrong, E. (2017). Twelve tips for applying the science of learning to health professions education. *Medical Teacher*, *39*, 26–31. doi: 10.1080/0142159X.2016.1231913

Horvath, A. O., & Greenberg, L. S. (1989). Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology*, *36*, 223–233. doi: 10.1037/0022-0167.36.2.223

Imel, Z. E., Baldwin, S. A., Baer, J. S., Hartzler, B., Dunn, C., Rosengren, D. B., & Atkins, D. C. (2014). Evaluating therapist adherence in motivational interviewing by comparing performance with standardized and real patients. *Journal of Consulting and Clinical Psychology*, *82*, 472–481. doi: 10.1037/a0036158

Ivers, N., Jamtvedt, G., Flottorp, S., Young, J. M., Odgaard-Jensen, J., French, S. D., … & Oxman, A. D. (2012). Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database of Systematic Reviews*, *6*, CD000259. doi: 10.1002/14651858.CD000259.pub3

Kerfoot, B. P., Baker, H. E., Koch, M. O., Connelly, D., Joseph, D. B., & Ritchey, M. L. (2007). Randomized, controlled trial of spaced education to urology residents in the United States and Canada. *Journal of Urology*, *177*, 1481–1487. doi: 10.1016/j.juro.2006.11.074

https://doi.org/10.1017/S1352465819000687 Published online by Cambridge University Press

La Rochelle, J. S., Durning, S. J., Pangaro, L. N., Artino, A. R., van der Vleuten, C. P., & Schuwirth, L. (2011). Authenticity of instruction and student performance: a prospective randomised trial. *Medical Education*, *45*, 807–817. doi: 10.1111/j.1365-2923.2011.03994.x

Ladany, N., Mori, Y., & Mehr, K. E. (2013). Effective and ineffective supervision. *The Counseling Psychologist*, *41*, 28–47. doi: 10.1177/0011000012442648

Lombardo, C., Milne, D., & Procter, R. (2009). Getting to the heart of clinical supervision: a theoretical review of the role of emotions in professional development. *Behavioural and Cognitive Psychotherapy*, *37*, 207–219. doi: 10.1017/S135246580900513X

Madson, M. B., Loignon, A. C., & Lane, C. (2009). Training in motivational interviewing: a systematic review. *Journal of Substance Abuse Treatment*, *36*, 101–109. doi: 10.1016/j.jsat.2008.05.005

Miller, W. R., & Moyers, T. B. (2017). Motivational interviewing and the clinical science of Carl Rogers. *Journal of Consulting and Clinical Psychology*, *85*, 757–766. doi: 10.1037/ccp0000179

Miller, W. R., & Rollnick, S. (2013). *Motivational Interviewing: Helping People Change*. New York, USA: Guilford Press.

Milne, D., & Reiser, R. P. (2012). A rationale for evidence-based clinical supervision. *Journal of Contemporary Psychotherapy*, *42*, 139–149. doi: 10.1007/s10879-011-9199-8

Moyers, T. B., Martin, T., Manuel, J. K., Hendrickson, S. M., & Miller, W. R. (2005). Assessing competence in the use of motivational interviewing. *Journal of Substance Abuse Treatment*, *28*, 19–26. doi: 10.1016/j.jsat.2004.11.001

Moyers, T. B., Martin, T., Manuel, J. K., Miller, W. R., & Ernst, D. (2010). Revised global scales: Motivational Interviewing Treatment Integrity 3.1.1 (MITI 3.1.1). Retrieved from: http://casaa.unm.edu/mimanuals.html

Moyers, T. B., Rowell, L. N., Manuel, J. K., Ernst, D. & Houck, J. M. (2016). The Motivational Interviewing Treatment Integrity Code (MITI 4): rationale, preliminary reliability and validity. *Journal of Substance Abuse Treatment*, *65*, 36–42. doi: 10.1016/j.jsat.2016.01.001

Parsons, M. B., Rollyson, J. H., & Reid, D. H. (2012). Evidence-based staff training: a guide for practitioners. *Behavior Analysis in Practice*, *5*, 2–11. Retrieved from: https://www.ncbi.nlm.nih.gov/pubmed/23730462

Rakovshik, S. G., & McManus, F. (2010). Establishing evidence-based training in cognitive behavioral therapy: a review of current empirical findings and theoretical guidance. *Clinical Psychology Review*, *30*, 496–516. doi: 10.1016/j.cpr.2010.03.004

Rakovshik, S. G., McManus, F., Vazquez-Montes, M., Muse, K., & Ougrin, D. (2016). Is supervision necessary? Examining the effects of internet-based CBT training with and without supervision. *Journal of Consulting and Clinical Psychology*, *84*, 191–199. doi: 10.1037/ccp0000079

Schwalbe, C. S., Oh, H. Y., & Zweben, A. (2014). Sustaining motivational interviewing: a meta-analysis of training studies. *Addiction*, *109*, 1287–1294. doi: 10.1111/add.12558

Simpson-Southward, C., Waller, G., & Hardy, G. E. (2016). Supervision for treatment of depression: an experimental study of the role of therapist gender and anxiety. *Behaviour Research and Therapy*, *77*, 17–22. doi: 10.1016/j.brat.2015.11.013

Tracey, T. J., & Kokotovic, A. M. (1989). Factor structure of the Working Alliance Inventory. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, *1*(3), 207–210. doi: 10.1037/1040-3590.1.3.207

Turner, L. B., Fischer, A. J., & Luiselli, J. K. (2016). Towards a competency-based, ethical, and socially valid approach to the supervision of applied behavior analytic trainees. *Behavior Analysis in Practice*, *9*, 287–298. doi: 10.1007/s40617-016-0121-4

van Merrienboer, J. J., & Sweller, J. (2010). Cognitive load theory in health professional education: design principles and strategies. *Medical Education*, *44*, 85–93. doi: 10.1111/j.1365-2923.2009.03498.x

Vervliet, B., Craske, M. G., & Hermans, D. (2013). Fear extinction and relapse: state of the art. *Annual Review of Clinical Psychology*, *9*, 215–248. doi: 10.1146/annurev-clinpsy-050212-185542

Watkins, C. E. (2011). Does psychotherapy supervision contribute to patient outcomes? Considering thirty years of research. *The Clinical Supervisor*, *30*, 235–256. doi: 10.1080/07325223.2011.619417

Wheeler, S., & Richards, K. (2007). The impact of clinical supervision on counsellors and therapists, their practice and their clients. A systematic review of the literature. *Counselling and Psychotherapy Research*, *7*, 54–65. doi: 10.1080/14733140601185274

Young, J. Q., Van Merrienboer, J., Durning, S., & Ten Cate, O. (2014). Cognitive load theory: implications for medical education: AMEE Guide No. 86. *Medical Teacher*, *36*, 371–384. doi: 10.3109/0142159X.2014.889290

Zarbock, G., Drews, M., Bodansky, A., & Dahme, B. (2009). The evaluation of supervision: construction of brief questionnaires for the supervisor and the supervisee. *Psychotherapy Research*, *19*, 194–204. doi: 10.1080/10503300802688478

Cite this article: Beckman M, Forsberg L, Lindqvist H, and Ghaderi A (2020). Providing objective feedback in supervision in motivational interviewing: results from a randomized controlled trial. *Behavioural and Cognitive Psychotherapy* **48**, 383–394. https://doi.org/10.1017/S1352465819000687

https://doi.org/10.1017/S1352465819000687 Published online by Cambridge University Press