Department of Clinical Neuroscience
Karolinska Institutet, Stockholm, Sweden

DISSEMINATION OF MOTIVATIONAL INTERVIEWING: THE ROLE OF WORKSHOP TRAINING AND SUBSEQUENT SUPERVISION IN THE DEVELOPMENT OF COMPETENCE IN CLINICAL PRACTICE

Maria Beckman

Karolinska Institutet

Stockholm 2018

Dissemination of Motivational Interviewing: The Role of Workshop Training and Subsequent Supervision in the Development of Competence in Clinical practice

THESIS FOR DOCTORAL DEGREE (Ph.D.)

Maria Beckman

Principal Supervisor
Professor Ata Ghaderi
Karolinska Institutet
Department of Clinical Neuroscience
Division of Psychology

Co-supervisors
Lars Forsberg
Margarita Diez

Opponent
Professor Steve Martino
Yale University
Department of Medicine
Division of Psychiatry

Examination Board
Professor Ulrica von Thiele Schwarz
Karolinska Institutet, Department of Learning, Informatics, Management and Ethics (LIME)
Mälardalens Högskola, Department of Psychology

Associate Professor Anette Søgaard Nielsen
Syddansk Universitet
Research Unit of Clinical Alcohol Research

Associate Professor Knut Sundell
The National Board of Health and Welfare
The Swedish Agency for Health Technology Assessment and Assessment of Social Services (SBU)

ABSTRACT

Aim: Training practitioners is often used as the primary implementation strategy to disseminate evidence based practices into clinical and community settings. The overall objective of this thesis was to evaluate various aspects of training practitioners as part of the dissemination and implementation of motivational interviewing (MI).

Method: In Study I, ten supervisors from a randomized controlled trial responded to semi-structured interviews about their supervision behaviors. Study II and III were conducted with practitioners in five Swedish county councils: In Study II, 174 practitioners were randomized to: 1) Regular county council workshop training; or 2) Regular county council workshop training followed by telephone supervision sessions based on objective feedback. All participants recorded three to eight sessions with actors. In Study III, the 98 participants from Study II receiving supervision were randomized to objective feedback based on either: 1) Half of a feedback protocol; or 2) The entire protocol. In study IV, 134 employees from The Swedish National Board of Institutional Care (SiS) with at least one completed MI training course were randomized to: 1) Six months of regular SiS supervision; or 2) Six monthly sessions of telephone supervision based on objective feedback. Study IV also replicated Study III. All participants in Study IV recorded three to seven sessions with a client or a colleague.

Results: While many of the reported supervisory behaviors in Study I were similar, there where also variations. Moreover, none of the supervisors described the feedback protocol as relatively important for the supervisees to learn MI, and half of them expressed concerns regarding an eventual negative impact of the objective feedback. In Study II, the different county councils workshops trainings increased the participants' MI skills to virtually the same level, and the subsequent supervision group showed larger proficiency gains at follow-up. In Study III, the group receiving feedback based on half of the protocol performed better at only two of the seven skill measures, and the objective feedback did not negatively affect the supervisory relationship or provoke supervisee discomfort/distress. In study III, many participants met the benchmarks for beginning proficiency already at baseline, and the regular group supervision and the supervision based on objective feedback were equally effective. The group that received feedback on half of the protocol performed better on only one of the seven skill measures, and the feedback did not negatively affect the supervisory relationship or provoke supervisee discomfort/distress.

Conclusions: In accordance with previous research, both Study II and IV showed that workshop training can increase participants' MI skills, and that subsequent supervision can further enhance acquired skills. Additionally, both studies indicate that these results also apply to naturalistic settings. However, the high variations in competence at all assessment points, together with the low interest in the possibility of subsequent supervision in both studies, are troublesome. In addition, neither the workshop trainings, nor the costly additional sessions of individual telephone supervision, or the comprehensive MI-implementation within SIS, were sufficient for many of the participants to reach beginning proficiency levels. This raises questions regarding both the most efficient form of training for practitioners to attain and sustain adequate practice standards, and how to create an interest among practitioners to participate in such training. Moreover, the results from Study III and IV showed that objective feedback does not seem to provoke significant supervisee anxiety or negatively affect the supervisory relationship. Although restricting the number of variables when providing objective feedback might promote learning during supervision, the observed differences in skill acquisition in both these studies were small, and it is not clear what really generated them. Since objective feedback seem to be an important part of supervision, and efficient supervision an important factor for the dissemination and implementation of evidence based practices, constructive replications are needed to ascertain the mode and complexity of feedback that optimizes practitioners' learning.

LIST OF SCIENTIFIC PAPERS

I. Beckman, M., Bohman, B., Forsberg, L., Rasmussen, F., & Ghaderi, A. (2017). Supervision in Motivational Interviewing: An Exploratory Study. *Behavioural and Cognitive Psychotherapy*, 45(4), 351-365. doi:10.1017/S135246581700011X

II. Beckman, M., Forsberg, L., Lindqvist, H., Diez, M., Eno Persson, J., & Ghaderi, A. (2017). The Dissemination of Motivational Interviewing in Swedish County Councils: Results of a Randomized Controlled Trial. *PloS One*, 12(7), e0181715. doi:10.1371/journal.pone.0181715

III. Beckman, M., Forsberg, L., Lindqvist, H., & Ghaderi, A. Providing Objective Feedback in Supervision in Motivational Interviewing: Results from a Randomized Controlled Trial. *Manuscript submitted for publication*.

IV. Beckman, M., Lindqvist, H., Forsberg, L., Lundgren, T., & Ghaderi, A. The Implementation of Motivational Interviewing in Institutions for Compulsory Care of Substance Abuse: Results of a Randomized Controlled Trial. *Manuscript submitted for publication*.

CONTENTS

LIST OF ABBREVIATIONS

CBS	Competency-Based Supervision
CBT	Cognitive Behavioral Therapy
CHC	Child Health Centers
CLT	Cognitive Load Theory
D&I	Dissemination and Implementation
EBCS	Evidence Based Clinical Supervision
EBT	Evidence-Based Treatment
EBP	Evidence-Based Practice
GLMM	Generalized Linear Mixed Model
ICC	Intraclass Correlation Coefficient
LVM	The Care of Substance Abusers Special Provisions Act
MI	Motivational Interviewing
MINT	The Motivational Interviewing Network of Trainers
MITI	The Motivational Interviewing Treatment Integrity Code
MIQA	The Motivational Interviewing Quality Assurance Group
SiS	The Swedish National Board of Institutional Care
SUD	Substance Use Disorder
WAI-S	The Short Working Alliance Inventory

1   INTRODUCTION

Each year, billions of dollars are spent globally in the public and private sectors on clinical and health services research, training healthcare practitioners, and quality improvement of healthcare (Grimshaw, Eccles, Lavis, Hill, & Squires, 2012). Despite this, a consistent finding is the failure to translate research to clinical practice (Grimshaw et al., 2012). As a result of this evidence-practice gaps, patients miss out on advances in healthcare, may not receive optimal care, and are sometimes exposed to potentially harmful treatments (Grimshaw et al., 2012).

Dissemination and implementation (D&I) science seeks to understand how to systematically facilitate deployment and utilization of evidence-based approaches to improve the quality and effectiveness of health care, and over 60 different models (i.e., theories and frameworks) for D&I have been described in the literature (Powell et al., 2012; Powell, Proctor, & Glass, 2014; Tabak, Khoong, Chambers, & Brownson, 2012). Dissemination has been defined as *the targeted distribution of information to a specific audience*, and implementation as *the use of strategies to introduce or adapt evidence-based interventions within specific settings* (Herschell, Kolko, Baumann, & Davis, 2010). Evidence-based treatment (EBT) has been defined as *interventions or techniques that have produced therapeutic change in controlled trials*, and evidence-based practice (EBP) as a broader term referring to *clinical practice that is informed by evidence about interventions, clinical expertise, and client needs, values, and preferences, and their integration in decision-making about individual care* (Kazdin, 2008).

Most of the implementation models describe the transfer of evidence to practice as a complex, multi-level process, using similar or overlapping constructs with slightly different terminologies and definitions (Damschroder et al., 2009; Fixsen, Naoom, Blase, Friedman, & Wallace, 2005; Kirk et al., 2016). However, training practitioners is still often used as the primary implementation strategy in the efforts to disseminate EBT into clinical and community settings (Edmunds, Beidas, & Kendall, 2013). Training practitioners as part of the D&I of motivational interviewing (MI) is also the focus of this dissertation.

1.1   MOTIVATIONAL INTERVIEWING

MI is a client-centered and directional therapeutic method for strengthening clients' motivation to change (Miller & Rollnick, 2013). MI was originally developed for treating substance abuse, but is now applied to a variety of clinical and healthcare settings (Moyers, Rowell, Manuel, Ernst, & Houck, 2016). In Sweden, extensive amount of resources have been invested in the implementation of MI in health care, social services, schools and correctional treatment. MI is also a key method in the national board of health and welfare's recommendations for evidence-based methods of preventing disease, and for treatments of substance use disorders http://www.socialstyrelsen.se/riktlinjer/nationellariktlinjer.

MI theory proposes a linguistic process wherein specific therapist technical verbal behaviors (i.e., MI-consistent and inconsistent skills) and relational factors (i.e., empathy and MI spirit) lead to client verbal behaviors (i.e., change talk and sustain talk), which in turn predict client

behavior change (Figure 1) (Pace et al., 2017). As a stand-alone treatment, MI usually consists of one to four sessions. MI is also often used as a pre-treatment or as an integral part of other treatments and interventions (Miller & Rollnick, 2013).
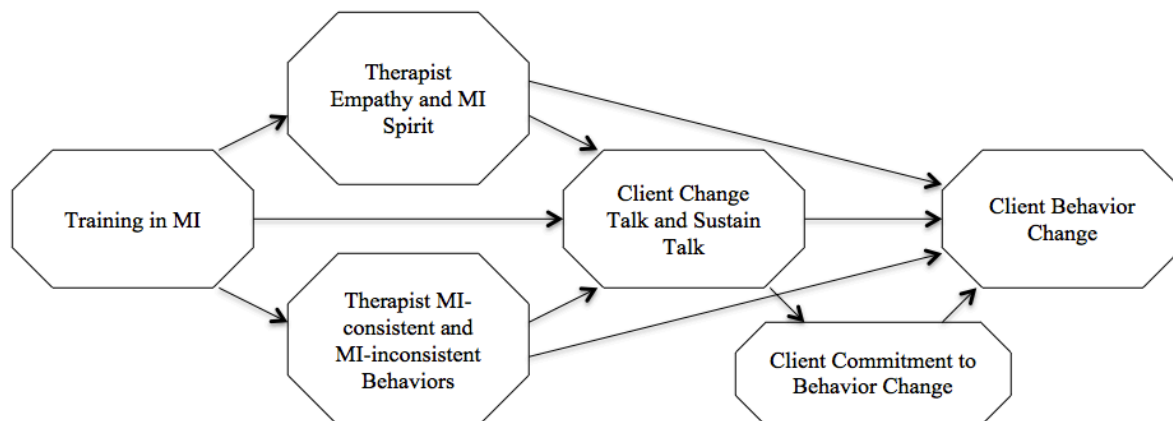

Figure 1. Hypothesized relationships among process and outcome variables in motivational interviewing (MI) (W. R. Miller & Rose, 2009).

Meta-analyses of MI have reported small to medium effects on average, with significant efficacy in relation to substance use, smoking, short-term weight loss, gambling, and certain improved medical outcomes (Miller & Rollnick, 2014). Still, the efficacy of MI, as that of all other counseling interventions, relies heavily on the adherence and competence of the practitioners practicing the method: Even if the support for many of the pathways of the technical and relational linguistic process model of MI still is inconclusive, two resent meta-analysis of MI process and outcome have found significant relationships between MI-consistent therapist behaviors and greater client change talk, as well as greater client sustain talk. The two meta-analysis also found correlations between MI-inconsistent therapist behaviors and more sustain talk, but not change talk, and when these indicators were combined into proportions, the overall technical hypothesis was supported. The relational hypothesis was not supported in any of the meta-analysis, but heterogeneity in the technical hypothesis path effect sizes was partially explained by inter- and intrapersonal moderators (Magill et al., 2018; Pace et al., 2017). Additionally, accurate empathy has shown positive client outcomes in research on other behavioral interventions (Elliott, Bohart, Watson, & Greenberg, 2011; T. B. Moyers & Miller, 2013).

Training practitioners to sustained levels of adherence and competence in MI has proven to be both costly and time-consuming (Hall, Staiger, Simpson, Best, & Lubman, 2015), and the format and amount of training required to achieve sustained practice change with impact on client outcomes are unclear (Hall et al., 2015). Previous research also indicates that the most common training format (i.e., a one-time workshop) may be insufficient for providing most practitioners with long term proficiency in MI, and that training tailored to practitioners'

context, together with subsequent supervision including systematic feedback, is more likely to produce enhanced and sustained skills (Barwick, Bennett, Johnson, McGowan, & Moore, 2012; de Roten, Zimmermann, Ortega, & Despland, 2013; Madson, Loignon, & Lane, 2009; Schwalbe, Oh, & Zweben, 2014; Soderlund, Madson, Rubak, & Nilsen, 2011). A recent meta-analysis of MI training studies found that approximately three to four post-workshop feedback and coaching sessions over a six-month period is sufficient to maintain workshop training effects (Schwalbe et al., 2014). Other researchers have instead proposed that ongoing training until proficiency is achieved makes more sense than relying on a fixed training dose (Hall et al., 2015; Martino, Canning-Ball, Carroll, & Rounsaville, 2011; Miller & Moyers, 2017; Miller & Rollnick, 2014). These conclusions regarding the difficulties in training practitioners to sustained levels of adherence and competence in MI are consistent with the conclusions from the broader literature on training practitioners.

1.2   TRAINING PRACTITIONERS

There are a number of educative strategies in the literature, including but not limited to: The distribution of educational materials, workshop trainings, the creation of practice manuals and guidelines, computer assisted decision-making, and workshop follow-ups or subsequent support strategies (e.g., consultation, supervision, audit and feedback, and peer support networks) (Edmunds et al., 2013).

A systematic literature review from a systems-contextual perspective of studies training therapists in EBP, ranging from 1990 to 2008, concluded that: 1) Therapists perceived and declarative knowledge increase following training; 2) Therapists attitudes improve after training and this is maintained at follow-up; 3) Self-reported therapist behavior change does not match actual behavior change; 4) Therapists trained in the most common format (i.e., workshop, manual, and brief supervision) do not reach proficiency in treatment adherence, competence, and skill; 5) There is insufficient information regarding how therapist variables, client characteristics, and organizational variables influence therapist behavior (e.g., adherence, competence, and skill) following training; 6) Therapist training in EBP does not currently engender improved client outcomes; 7) The quality of training is important to engender client change: Active learning during training is integral to influence both therapist and client change; and 8) When addressing therapist and client variables, as well as organizational support and training, therapists may reach sufficient proficiency levels to influence client change (Beidas & Kendall, 2010).

Another empirical review of 55 psychotherapy training studies concluded that multi-component training have been studied most often and have most consistently demonstrated positive training outcomes relative to other training methods, that workshop follow-ups help to sustain outcomes, and that little is known about the impact of train-the-trainer methods (Herschell et al., 2010).

Rakovshik and McManus (2010) review of the efficacy and effectiveness of training in 37 cognitive behavioral therapy (CBT) training, effectiveness and dissemination studies concluded that: 1) Traditional dissemination approaches (i.e., workshops and manuals) seem insufficient in producing significant change in both therapists' skills and clients' outcomes; 2)

Training variably leads to increased therapist competence, which is positively related to better client outcome; 3) Extensive training is associated with improvement in therapists' competence and client outcome; 4) Graded training (i.e., extended instruction for therapists with a more tardy learning curve) allows for efficient allocation of resources in heterogeneous training groups; 5) Some mode of theoretical instruction seems integral in initial training. This may be provided through reading, web-based instructions or in workshops, as long as it is followed by experiential and interactive training through practice cases, co-therapy or supervision; 6) Adherence monitoring followed by feedback and instructions may be a productive and necessary focus in training; 7) Sustained supervision over a prolonged period may be necessary to maintain competency gains; and 8) Broader constructs such as dosage, spacing, sequence and scaffolding indicate the pertinence of theoretical models of learning (Rakovshik & McManus, 2010).

One additional review of 27 randomized studies examining the efficacy of CBT interventions for depression and anxiety disorders concluded that the ubiquity of intervention-specific training in research contexts risks being overlooked when commissioning evidence-based therapies in routine practice, with clear implications for the likely effectiveness of the interventions (A. D. Roth, Pilling, & Turner, 2010).

The most definitive conclusion from all this research on training practitioners is that brief workshops can influence therapist knowledge, attitudes, and perceived behavior, but in order to change practitioners' actual behavior, active, behaviorally-oriented training techniques (e.g., objective feedback, reflective practice, and behavioral rehearsal/role-play) and subsequent supervision should be included (Beidas, Edmunds, Marcus, & Kendall, 2012; Beidas & Kendall, 2010; Edmunds et al., 2013; Herschell et al., 2010; Rakovshik & McManus, 2010; A. D. Roth et al., 2010; Webster-Stratton, Reid, & Marsenich, 2014). These training strategies are also in line with the standard methods used to train practitioners in high-quality clinical trials to ensure that they adhere to treatment protocols and deliver therapy in a competent manner (A. D. Roth et al., 2010; Sholomskas et al., 2005). However, despite the increased attention to supervision as part of training programs for practitioners, supervisory practice still lacks a well-developed empirical base and there is uncertainty regarding exactly how it should be provided (Reiser & Milne, 2012).

1.3 SUPERVISION

Supervision has been defined in a number of ways and can be referred to a variety of settings. Common to most definitions is that a more experienced professional works with a less experienced member of the same profession in an evaluative relationship that extends over time (Bernard & Goodyear, 2014). Even if research on supervision is sparse (Beidas & Kendall, 2010; Herschell et al., 2010), a number of supervisory components can be discerned in the literature; supervisors who are available, supportive, function as a professional role model, and delegate responsibility to the supervisees are highly valued by their supervisees (Bogo & McKnight, 2006). Highly valued are also evaluation practices, regular feedback, and supervisors who are knowledgeable about tasks and skills and can relate these to theory (Bogo & McKnight, 2006; C. A. Falender & Shafranske, 2014).

Research has moreover found that supervision can have a beneficial effect on supervisees with an impact on self-awareness, skills, self-efficacy, theoretical orientation, support, and the supervisee-client relationship (C. E. Watkins, 2011; Wheeler & Richards, 2007). There is also increasing evidence that inclusion of supervision in training programs for practitioners is associated with better outcomes for both trainees and clients (Rakovshik, McManus, Vazquez-Montes, Muse, & Ougrin, 2016). However, the methodological shortcomings of this research make it challenging to identify the particular aspects of supervision that comprise best practices (Bearman, Schneiderman, & Zoloth, 2016).

The history of supervision stretches back over 100 years. An informal start of psychoanalytic supervision often mentioned is Freud's case discussion meetings held at his home in the early 20th century, while other approaches to psychotherapy supervision began to form around the 1950s (C. E. Watkins, 2012a). The current basis of CBT supervision largely rests on descriptions provided by Liese and Beck (1997) and Padesky (1996). In both these texts, the structure of supervision incorporates activities analogous to those that characterize a typical CBT session: Checking in, bridging back to the last session, collaborative agenda setting, didactic instructions, discussions, modeling, behavioral rehearsal/role-plays, summarizing the session, assigning homework, and receiving feedback about the session from the supervisee (Reiser & Milne, 2012).

In line with the psychotherapy approach, CBT supervision focuses on specific training goals together with plans and strategies for attaining them (Johnson & Kaslow, 2014). These CBT supervision processes aims to facilitate the supervisees' assimilation of the clinical frame, promote proficiency in the application of techniques, and encourage clinical creativity (Rosenbaum & Ronen, 1998). Friedberg and colleagues (2010) have conceptualized three broad competency domains for CBT supervision: (1) Declarative knowledge (i.e., factual knowledge of theory, research, and technique); (2) Procedural knowledge (i.e., knowledge that involves implicit learning exercised in the performance of skills); and (3) Self-reflective capacities (i.e., the capacity to monitor, understand, and evaluate one's own actions and their clinical impact).

Newman (2010) has defined some fundamental competency domains in an account of contributions from CBT supervision to competency-based professional psychology training: (1) Competency in utilization and evaluation of scientific research; (2) Relationships competency; (3) Individual and cultural diversity competency; and (4) Competency in working within interdisciplinary systems and interdisciplinary health care teams. However, the CBT supervision literature is largely descriptive and emphasizes on principles (e.g., structure, relationship factors, collaboration and guided discovery), rather than on explicit procedures or manuals. Also, while the practice of CBT has been supported by a tradition of empirical science, just as other supervision practices, CBT supervision lacks a convincing empirical base as well as well-defined supervisory procedures that can be reliably observed and measured (Reiser & Milne, 2012).

Some researchers have tried to clarify the differences between ongoing supervision, where practitioners are supervised as part of their work within an organization, and expert external support following training as a D&I implementation strategy, by using the term consultation

for the latter (Dorsey et al., 2018; Edmunds et al., 2013; Nadeem, Gleacher, & Beidas, 2013; Sewell, 2017). In a recent study of verbal interactions from recordings of 438 ongoing workplace-based supervision sessions (Dorsey et al., 2018), supervision techniques most often used to train counselors within research studies (i.e., direct observation of treatment sessions, behavioral rehearsal/role-plays and modeling), were used only rarely, which led the authors to suggest that ongoing workplace-based supervision can provide strategies to support implementation of EBP, but to also suggest areas for improvement (Dorsey et al., 2018).

Due to this lack of well-defined supervisory procedures grounded in an empirical base in general supervision practice, a movement towards more evidence and competency based supervision practice has emerged in recent years (E. Watkins, 2014). A growing number of comprehensive frameworks for supervision practices can, among other things, be seen as an expression of this movement (American Psychological Association, 2014; Borders et al., 2011; A. Roth & Pilling, 2009; The Psychology Board of Australia, 2013). While each framework has some distinctive elements, they are all highly similar in their intent and content; providing specifics about the knowledge, abilities, skills and attitudes that are expected for adequate supervision practice (E. Watkins, 2014).

Another expression of this evidence and competency based supervision movement is the emerging approaches to supervision that cuts across theoretical lines. Competency-based supervision (CBS) is a meta-theoretical approach that does not preclude other models of supervision in its impetus to change the viewpoints of training in supervision, defined as *an approach that explicitly identifies the knowledge, skills, and values that are assembled to form a clinical competency and develops learning strategies and evaluation procedures to meet criterion referenced competence standards in keeping with evidence-based practices and requirements of the local clinical setting* (C. A. Falender & Shafranske, 2014). CBS entails a focus on adjusting supervision to the supervisee's current functioning, monitoring the supervisee's development, assessing the supervisee's competencies, and ensuring accurate and timely feedback to the supervisee.

Evidence based clinical supervision (EBCS) is another emerging approach to supervision that cuts across theoretical lines (C. E. Watkins, 2012b). EBCS is characterized by supervision practices based on the best evidence available, that is, those specific supervision practices that are supported by evidence and systematic analysis of supervision efficacy (Milne & Reiser, 2012). An evidence-based supervision framework has the distinct objective to facilitate the full and balanced experiential learning of the supervisee (Milne & Reiser, 2012), and has the potential to: (1) More effectively ensue and enhance supervision accountability and progress over time; (2) Improve and guide supervision development; (3) Aid and inform clinical decision making; (4) Monitor and protect client care; and (5) Support fidelity in supervision implementation (C. E. Watkins, 2012b).

Another evidence-based approach to supervision is the specific clinical trials supervision most often used to train counselors within research studies (Martino, Gallon, Ball, & Carroll, 2008). Most clinical trials use a similar approach to train counselors to deliver treatments with fidelity, and according to Martino and colleagues (2008) this approach could also be

used in regular supervision practice. The three core supervision components of this approach are: (1) Direct observation of treatment sessions; (2) Structured performance feedback on counselors' treatment adherence and competence; and (3) Coaching to improve the ability of counselors to implement treatments proficiently. In a clinical trials' training approach to supervision, the supervisor's performance feedback is structured, always based on counselors' recordings of sessions and includes performance strengths and weaknesses together with opportunities for counselor self-evaluation and input.

All these supervision practices share some core components: Articulations of supervision goals, assessment feedback based on monitoring of sessions, individualized coaching/training including active learning (e.g., performance feedback, reflection and behavioral rehearsal/role-plays), and evaluation of the targeted supervision competencies. However, despite the increased attention to objective feedback as part of efficient supervision, issues of how to best deliver feedback to promote learning remain a matter of ongoing debate (Bosse et al., 2015).

1.4   OBJECTIVE FEEDBACK

Performance-related feedback is perceived as an important part of clinical supervision and has demonstrated small but potentially important improvements in clinical practice (Ivers et al., 2012). Yet, it is still unclear how or how much feedback should be delivered to maximize its effect, as well as how feedback and learners' activity are optimally blended (Bosse et al., 2015; Ivers et al., 2012). Feedback typically contains evaluation of supervisee's skills and adherence (Parsons, Rollyson, & Reid, 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 behaviors, and may also aid the supervisee to self-assess more accurately (Parsons et al., 2012).

Some previous research propose that supervisees may fear negative evaluations from their supervisors (Abernethy & Cook, 2011; Bernard & Goodyear, 2014; Clarke & Giordano, 2013; Ellis, Hutman, & Chapin, 2015; Friedberg, Gorman, & Beidel, 2009; Lombardo, Milne, & Procter, 2009), and that supervisors often feel critical and worry that their feedback may harm the supervisory working alliance (Chur-Hansen & 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, Fischer, & Luiselli, 2016). Previous research has also found providing feedback as one of the weaker skill areas for clinical supervisors, a difficulty that might serve to generate fear of evaluation in the supervisee (Cummings, Ballantyne, & Scallion, 2015).

High levels of supervisee anxiety have been suggested to trigger supervisee defensiveness, reduce supervisees' willingness to disclose information, cause supervisor–supervisee role conflicts, and decrease supervisee clinical performance (Ellis et al., 2015). Findings from the science of learning also points to emotions as an important aspect to consider when training practitioners: Emotions affect attention to information as well as how memory is stored and retrieved (Gooding, Mann, & Armstrong, 2017). Moreover, positive emotions during learning have been associated with deeper cognitive processing and enhanced learning, while negative

emotions (i.e. anxiety) have been associated with more superficial cognitive processing and impeded learning (Young, Van Merrienboer, Durning, & Ten Cate, 2014).

Other findings from the supervision field suggest that the common assumption of recordings and monitoring of sessions as overburdening for the supervisee is a misconception, and that supervisees in general both quickly adopt to recordings of sessions and can handle corrective feedback, especially within a positive supervisory working alliance (Ellis, 2010; Ladany, Mori, & Mehr, 2013). Previous research also concludes that supervisees often complain about vague corrective, or exclusively positive, feedback, and that supervisees are more satisfied when supervisors provide both positive and corrective feedback (Chur-Hansen & McLean, 2006; Cummings et al., 2015). Some researcher even suggest that providing both positive and corrective feedback enhances the supervisory relationship and increases the supervisors satisfaction with their role (Chur-Hansen & McLean, 2006).

The Motivational Interviewing Treatment Integrity (MITI) Code is a coding system with acceptable psychometric properties (Forsberg, Kallmen, Hermansson, Berman, & Helgason, 2007; Moyers, Martin, Manuel, Hendrickson, & Miller, 2005), widely used as a treatment integrity measure and as a feedback tool in MI training and supervision. MITI 3.1 (Moyers, Martin, Manuel, Miller, & Ernst, 2010) consists of two main components: (1) The five global dimensions (Empathy, Evocation, Collaboration, Autonomy and Direction) that reflects the coder's overall judgment of the practitioner's performance on a 5-point scale; and (2) The behavior counts, which are frequency counts of the practitioners' of every utterance coded in seven specified categories (Giving information, MI adherent behaviors, MI non-adherent behaviors, Closed questions, Open questions, Simple reflections and Complex reflections).

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 clinical skills. However, although many researchers emphasize feedback as crucial for clinical learning (Bosse et al., 2015), several questions remain regarding how to best deliver feedback to promote the acquisition of skills, especially when using multifaceted feedback tools as the MITI, to promote learning during supervision.

Several meta-analyses on learning have established that the average effects attributable to feedback are among the highest we know in education, but also that feedback effects are among the most variable in their influences (Benassi, Overson, & Hakala, 2014). These variances indicates that some types and circumstances of feedback are more powerful than others, and that merely prescribing high dosages of feedback does not imply that learning will take place: "It seems we know much about the power of feedback, but too little about how to harness this power and make it work more effectively" (Benassi et al., 2014). One effect of feedback appears to be in influencing the effort students allocate to an activity: Student seem to devote more time and effort to tasks where specific and timely feedback are available (Benassi et al., 2014). Other aspects from the science of learning could also be important to consider when planning training programs with subsequent supervision as part of the D&I of EBT.

1.5   RESEARCH ON LEARNING

Successful learning requires an interplay of multiple processes, including those in the cognitive, affective, social, environmental and metacognitive domains, and over the years, many often overlapping theories of learning have been put forward (Young et al., 2014).

Cognitive Load Theory (CLT) is one of the leading cognitive learning theories, with implications for both complex learning and skills acquisition (van Merrienboer & Sweller, 2010). 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 situation: 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, according to the theory, be reduced or even hindered (van Merrienboer & Sweller, 2010).

The CLT framework proposes educational design strategies used to reduce cognitive load that has received increased recognition in recent medical education (Young et al., 2014). Starting with fewer elements, reducing irrelevant material, chunking elements into more manageable parts, or highlighting essential material, are all examples of such strategies to free up space for processing complex information (Gooding et al., 2017). Other studies within the CLT framework have shown that also initial level of knowledge can affect the efficacy of instructional methods: Designs and techniques effective with unskilled learners can be less effective and even have negative consequences for more proficient learners (Kalyuga, 2007). This reversal in the relative effectiveness of instructional methods with increased levels of skills has been referred as "the expertise reversal effect" (Kalyuga, 2007). In addition to the CLT framework, other research in the cognitive sciences on memory and learning have resulted in a set of generic learning principles applicable to a wide range of learning situations and settings (Cutting & Saks, 2012). Examples of such generic learning principles relevant to training practitioners are:

1) *Spaced practice* – Spaced practice is a strategy in which the learner is exposed to educational material on at least two occasions, separated by a period of time. Research has shown that, if a given amount of study time is distributed or spaced across multiple sessions, as opposed to massed into a single session, performance is superior (Benassi et al., 2014; Pashler, Rohrer, Cepeda, & Carpenter, 2007; Rohrer & Pashler, 2010). The optimal study gap seem to increase with the duration of the test delay, and a longer-than optimal spacing does not seem to be nearly as harmful as a shorter-than-optimal spacing (Rohrer & Pashler, 2010). Some data suggest that the interval should be 5-10% of the time the information should be retained (i.e., monthly if the goal is to remember the information at one year) (Gooding et al., 2017). If the goal is a lifelong retention – the aim in most educational contexts – then previously studied material should be revisited at least a year after the first exposure (Rohrer & Pashler, 2010). The spacing effect appears to hold over educationally relevant time periods for both simpler cognitive tasks and more complex forms of learning, such as mathematics and surgical performance (Rohrer & Pashler, 2010).

2) *Interleaving* – Interleaving is a strategy where the learners mix, or interleave, multiple subjects or topics (e.g., abcbcacab), as opposed to blocked practice where the learner study one topic thoroughly before moving to the next (e.g., aaabbbccc) (Benassi et al., 2014; Rohrer & Pashler, 2010). Interleaving has shown to be more effective than blocked practice for long-term retention and improved ability to transfer learned knowledge, as well as for developing both motor skill learning and cognitive skills such as discriminability, categorization and problem solving (Rohrer & Pashler, 2010). Interleaving is sometimes also referred to as cumulative review.

3) *Desirable difficulties* – Research on durable learning has found that some procedures that produce more errors and are perceived as more difficult during the learning session can be beneficial for long-term learning (Cutting & Saks, 2012; Gooding et al., 2017; Rohrer & Pashler, 2010). These procedures have been called "desirable difficulties" (Schmidt & Bjork, 2012), and should not be confused with cognitive load. Instead, desirable levels of difficulty implies that the learning experience should be sufficiently difficult to promote mental effort, but not so difficult that learners cannot engage appropriately with the material (Cutting & Saks, 2012).

4) *Retrieval practice* – Retrieval practice is a learning strategy that requires the learner to recall information previously acquired. Research shows that retrieval practice both enhances learning and slows the rate of forgetting (Cutting & Saks, 2012; Gooding et al., 2017; Pashler et al., 2007; Rohrer & Pashler, 2010). Generative retrieval, where learners generate their own answers, leads to more durable learning than choosing answers from multiple-choice question, and providing feedback on the retrieval enhances learning further (Gooding et al., 2017). The type of retrieval practice should match the complexity of the information or task to be remembered – When learning facts, simple factual recall is often sufficient, but when learning how to perform a more complex task, practicing the whole task at regular intervals is necessary (Gooding et al., 2017). Retrieval practice is sometimes also referred to as the testing effect or test-enhanced learning.

5) *Metacognition* – Metacognition, or thinking about thinking, is the set of processes involved in monitoring one's own thinking (Benassi et al., 2014). Metacognition is critical for learners' ability to direct their ongoing learning techniques (Gooding et al., 2017), and learners who employ metacognitive strategies are able to better evaluate their progress and make decisions about strategies for improvement (Cutting & Saks, 2012).

6) *Transfer of learning* – The ability to apply acquired skills in a new and different context is known as transfer (Gooding et al., 2017; Young et al., 2014). Transfer of learning does not automatically occur and is therefore something that needs to be developed by learning examples designed around realistic situations, whole-tasks in increasingly realistic settings, or by providing the learners relevant and meaningful contexts for learning (Young et al., 2014).

All these findings from the science of learning (i.e., CLT, the expertise reversal effect, spaced practice, interleaving or cumulative review, desirable difficulties, retrieval practice or testing effect/test-enhanced learning, metacognition, and transfer of learning) have clear implications for training programs with subsequent supervision as part of the D&I of EBT.

2 OBJECTIVES

Training practitioners is often used as the primary implementation strategy in the efforts to disseminate EBT into clinical and community settings. The overall objective of the thesis was to evaluate various aspects of training of practitioners as part of the D&I of MI. Specific aims for each study are presented below.

2.1 STUDY I

The aim of the first study was to explore ten MI supervisors' behavior in a primary prevention intervention of childhood obesity conducted at child health centres in Sweden, in order to identify factors that might facilitate supervisees' proficiency in MI by contrasting them with current models of effective supervision.

2.2 STUDY II

The aim of the second study was to evaluate to what extent practitioners offered workshop trainings in MI through community-based implementation programs in Sweden acquire and retain skills from regular workshop trainings, as opposed to workshop trainings followed by supervision consisting of feedback based on monitoring of practice.

2.3 STUDY III

The aim of the third study was to evaluate how two different ways to provide objective feedback during MI supervision affects the supervisees' MI skill acquisition, perceived feelings of discomfort or distress provoked by the supervision sessions, and the supervisor-supervisee working alliance.

2.4 STUDY IV

The aim of the fourth study was to assess the MI skills of practitioners within The Swedish National Board of Institutional Care (SIS) with previous training in MI, and to evaluate the effectiveness of different ways to provide subsequent supervision on the supervisees' MI skill acquisition. Additionally, the fourth study also replicated Study III.

3 METHODS

An overview of the studies included in this thesis is presented in Table 1. Study I had an exploratory study design, Studies II, III, and IV had a confirmatory study design.

Table 1. Design, participants, data collection, and statistical analyses of the included studies.

Study	Design	Participants	Data collection	Statistical analyses
I	Descriptive Exploratory Study	MI supervisors ($n = 10$)	Semi-structured interviews	Mixed method approach
II	Randomized Controlled Trial	Swedish county council practitioners ($n = 174$)	Questionnaires, MITI-coded recorded sessions	Generalized linear mixed model Chi-square analyses
III	Randomized Controlled Trial	Participants from Study II randomized to MI supervision ($n = 98$)	Questionnaires, MITI-coded recorded sessions	Generalized linear mixed model Chi-square analyses
IV	Randomized Controlled Trial	Practitioners from SiS ($n = 134$)	Questionnaires, MITI-coded recorded sessions	Generalized linear mixed model Chi-square analyses

Note. MITI = motivational interviewing treatment integrity code, SiS = The Swedish National Board of Institutional Care.

3.1 ETHICAL CONSIDERATIONS

All procedures that contribute to the work of the present thesis comply with the ethical standards of the Regional Ethical Review Board in Stockholm, Sweden, and with the Helsinki Declaration of 1975 and its most recent revision. All studies included have been approved by the Regional Ethical Review Board in Stockholm, Sweden (2006/525-31/2; 2012/2195-31/5; 2013/904-31).

3.2 STUDY SETTING

Study I was conducted with supervisors in the PRIMROSE trial, a randomized controlled primary prevention intervention of childhood obesity conducted in eight county councils of Sweden (Doring et al., 2014). Nurses at child health centers (CHC) conducted the manual-based MI program with participating families. Prior to the trial, the nurses received MI-training with subsequent supervision including objective MITI feedback.

Study II and III were conducted in five Swedish county councils that participated in an MI training program evaluation in the county councils, funded by the National Board of Health and Welfare. The workshop trainings were open for all employees in each county council to attend, after approval from respective healthcare providers. The participants thus came from a variety of publicly financed healthcare facilities, with different strategies for the D&I of MI in each facility. Additionally, the form and content of the county councils' workshops also differed slightly, with most variation in length (i.e., ranging from two to three and a half days).

Study IV was conducted in twelve SiS residential care facilities for adults taken into care under the Swedish Care of Substance Abusers Special Provisions Act (LVM). The MI-implementation within SIS is comprehensive: A project manager, together with four national trainers, all members of the Motivational Interviewing Network of Trainers (MINT), administers the work. Their main task is to organize and support the implementation, and to ensure the MI quality assurance within the organization. The four national trainers, together with eight additional MI-trainers – all members of MINT, also organized workshop trainings for all employees who interact with clients, extended trainings, and trainings for the SiS MI coaches. Moreover, another MI implementation team works with detecting synergies and delimitations from other methods and SiS government assignments, and to anchor MI within all parts of the organization

3.3   PARTICIPANTS AND PROCEDURES

3.3.1   Study 1
In the first study, ten MI supervisors responded to one-to-one semi-structured telephone interviews about their supervision behaviors. The mean age was 49.4 years ($SD = 9.9$, range = 34.0–68.0). Seven were females and nine had previous training in CBT. Five were psychologists, three were nurses, one had a university degree in behavioral sciences, and one was an occupational therapist. On average, the supervisors had 7.0 years ($SD = 6.5$, range = 2.5–25.0) of experience as MI trainers, and 4.0 years ($SD = 3.7$, range = 1.0–15.0) of experience as MI supervisors. All supervisors were members of the Motivational Interviewing Network of Trainers (MINT), which organizes practitioners trained to train others in MI. Participation was voluntary and confidential, and prior to the interviews all the supervisors signed an informed consent form. As there was no existing tool in the literature, the study authors developed the interview protocol. The protocol was sent by e-mail prior to the interview so the supervisors would have time to reflect on the questions. As the first author was one of the supervisors in the trial, the second author interviewed the first author. Thereafter, the first author interviewed all the other supervisors. The interviews lasted between 28 and 34 minutes.

3.3.2   Study II
In the second study, 174 county councils practitioners were randomized to either: 1) Regular county council workshop training; or 2) Regular county council workshop training followed by six individual sessions of telephone supervision based on objective feedback (i.e., the

MITI). The mean age of the participants was 43.3 years ($SD = 13.6$), and the majority were females ($n = 111$, 88.1%). The education level varied from bachelor's degree ($n = 78$, 61.9%) to master's degree ($n = 45$, 35.7%). Two participants chose not to disclose their education level. The participants had a variety of occupations: nurses ($n = 34$), clinicians in social services ($n = 18$), physiotherapists ($n = 18$), teachers ($n = 11$), medical doctors ($n = 10$), counselors ($n = 7$), psychologists ($n = 6$), dietitians ($n = 6$), occupational therapists ($n = 4$), assistant nurses ($n = 3$), audiologists ($n = 2$), career counselor ($n = 1$), coach ($n = 1$), dental hygienist ($n = 1$), interpreter ($n = 1$), podiatrist ($n = 1$), and speech-language pathologist ($n = 1$). One participant chose not to disclose a profession. All the participants in the study recorded three 20-minutes sessions over phone with one of five actors role-playing standardized patients: One before the county councils' workshop trainings (baseline), one directly after the workshop trainings (post-workshop), and one 6 months after (follow-up). The participants randomized to receive subsequent supervision recorded five additional sessions with the actors between the post-workshop and the follow-up recording.

3.3.3 Study III

In the third study, the 98 participants from Study II randomized to receive subsequent supervision were randomized to either: 1) Systematic feedback based on half of the MITI (i.e., the behavioral component only); or 2) Systematic feedback based on the entire protocol. In the sessions based on half of the MITI, the supervisor could speak in general terms about the global dimensions. However, the global scores had been removed from the protocol, and the supervisors were not allowed to mention anything regarding the supervisees' results on these scores.

3.3.4 Study IV

In the forth study, 134 SiS employees with at least one completed training course in MI were randomized to: 1) Six months of continued regular SiS group supervision; or 2) Six monthly sessions of individual telephone supervision based on objective feedback (i.e., the MITI). As a replication of Study III, the 70 participants in the telephone supervision group were then randomized to: 1) Supervision including systematic feedback based on only the behavior counts of the MITI; or 2) Supervision including systematic feedback based on the entire MITI. The mean age of the participants was 43.2 years ($SD = 10.2$), and the majority were females ($n = 84$, 62.7%). Their education level varied from college/higher degree ($n = 45$, 33.6%) to no higher degree ($n = 53$, 39.6%) (26.8% missing), and they had a variety of occupations: Head of institution ($n = 4$), treatment responsible ($n = 2$), treatment assistant ($n = 89$), treatment secretary ($n = 14$), nurse ($n = 3$), assistant nurses ($n = 3$), other ($n = 3$) (11.9% missing). All the participants in the study recorded three 20-minutes sessions with either a client or a colleague: At baseline, six months after the baseline recording, and twelve months after the baseline recording. The group randomized to subsequent supervision recorded four additional sessions with either a client or a colleague in between the baseline recording and the six-month assessment. The regular SiS group supervision sessions had various content at the different institutions. The individual telephone supervision sessions had the same form and content as the supervision sessions in Study II and III.

3.3.5 The Individual Telephone Supervision Sessions

The monthly individual telephone supervision sessions conducted in Study II, III and IV were 30 minutes long and performed by trained coders at the Motivational Interviewing Quality Assurance (MIQA) group at Karolinska institutet in Stockholm, Sweden. For each session, the supervisor both coded the recordings and preformed the supervision. Before each session, the participants were asked to listen through the recordings, and ten minutes before the session started, the supervisor emailed the protocol to the participant.

All the supervision sessions were based on the results of the MITI, conducted in a manner consistent with MI and structured by a manual (Appendix 1): The sessions started with an introduction and a collaboratively agreed upon agenda. The supervisor also reviewed homework from previous sessions during this initial segment. Then, referring to the results of the MITI, the supervisor provided performance feedback and initiated a discussion about consistent and inconsistent MI behaviors and gave opportunities for participant's self-evaluation and input. The greatest focus was on the practice phase, where the supervisors used individualized role-plays based on segments from the recorded sessions and demonstrated specific skills to promote learning through observation. The supervisors then encouraged the participants to practice one or two specific skills at home. At the end of each session, participants summarized and evaluated the session. The supervisors had supervisory meetings once a month throughout the study period, to which they brought a self-selected recording of a supervision session.

3.4 ASSESSMENT

3.4.1 The Motivational Interviewing Treatment Integrity Code

In Study II, III and IV, all recorded 20 minutes sessions were assessed for proficiency in MI with the Swedish version of the MITI, version 3.1 (Forsberg, Forsberg, Forsberg, van Loo, & Rönnqvist, 2011) by the coders at the MIQA group. The coders were not blind to the participants' group allocation. To aid the evaluation of clinicians' skillfulness in MI, MITI provides recommended indicators of MITI Beginning proficiency and Competency (Table 2).

To reach and maintain MITI inter-rater reliability, all the coders at MIQA have completed 120 hours of training. The coders at MIQA also participate in group-coding sessions every week, and twice a year, 12 randomly selected recordings are double-coded by all the coders to assess rating consistency. According to Cicchetti's (1994) system for evaluating intraclass correlations, an intraclass correlation coefficient (ICC) below 0.40 is considered poor, an ICC between $0.40 – 0.59$ is considered fair, an ICC between $0.60 – 0.74$ is considered good, and an ICC between $0.75 – 1.00$ is considered excellent (Cicchetti & Sparrow, 1981). Table 3 shows the MIQA coder's inter-rater reliability for Study II, III and IV, assessed in the middle of both of the study periods (January 2014 for Study II and III; January 2015 for Study IV).

Table 2. 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
Percent open questions (%OC)[3]	50%	70%
Percent complex reflections (%CR)[4]	40%	50%
Percent MI-adherent (% MIA)[5]	90%	100%

Note. 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 behaviors / (MI adherent + MI non-adherent behaviors)

3.4.2 The Interview Protocol

The study authors developed the interview protocol used in Study I, which consisted of 32 questions (Appendix 2). The first eight questions covered aspects of participants' demographics (gender, age, education, and professional experience). The remaining 24 questions covered PRIMROSE trial-specific aspects of MI supervision, as well as generic aspects of supervision based on research on supervision. Several response formats were used, including forced-choice, free text, and visual analogue scales.

3.4.3 The Questionnaires

In Study II, III and IV, the participant's characteristics (i.e., gender, age, education level and profession) were assessed with self-reported questionnaires. Additionally, in all these studies, the supervisees' experience of the supervision sessions was measured with an adapted version of the short Working Alliance Inventory (WAI-S) (Tracey & Kokotovic, 1989), with an additional question regarding discomfort/distress. WAI-S contains 12 items rated on a seven points scale. It provides a summary of the total working alliance and three subscales to assess primary components of the working alliance: Goal (i.e., agreement with regard to the treatment goals), task (i.e., agreement with regard to the tasks) and bond (i.e., the empathic bond between the client and the therapist). The adaptation consisted of minor reframing of items to apply to the relationship between the supervisee and the supervisor instead of client-therapist.

Eventual feelings of discomfort/distress provoked by the supervision sessions were assessed with one additional question at the end of the WAI-S questionnaire. The participants were asked to specify a number from one to ten that reflected their level of discomfort during the supervision sessions: 1 = *No feeling of discomfort/distress or anything that may fall under the*

category of a negative feeling or experience, 10 = *A feeling of discomfort/distress or a negative feeling or experience of any kind*.

Table 3. The MIQA coder's inter-rater reliability for Study II, III and IV, assessed with a two-way mixed model with absolute agreement, single measures, ICC.

MITI variable	ICC Study II & III	ICC Study IV
Empathy	.60	.69
Evocation	.69	.44
Collaboration	.74	.71
Autonomy	.75	.79
Direction	.49	.44
Giving information	.89	.62
MI adherent behaviors	.81	.82
MI non-adherent behaviors	.59	.58
Closed questions	.64	.97
Open questions	.92	.98
Simple reflections	.73	.82
Complex reflections	.68	.75

Note. MITI = motivational interviewing treatment integrity code, ICC = intraclass correlation coefficient.

3.5 STATISTICAL ANALYSES

All the analyses in the thesis were performed using the SPSS (Version 22, SPSS Inc., Chicago, IL). In Study I, both qualitative and quantitative data were collected and a mixed method approach was performed. For the qualitative analysis of the textual data from the interviews, a conventional qualitative content analysis (Hsieh & Shannon, 2005) was used. Coding was developed inductively (from the data content) as opposed to employing a predefined coding scheme. Following the transcription of the interviews, two members of the research team independently analyzed the supervisors' qualitative statements. Common themes and contrast statements (either words, phrases, or sentences) across interviews were extracted; data were coded and similar elements were summarized thematically. The two research team members then compared their results and discussed the ratings they disagreed upon until they reached a consensus.

For the quantitative analysis, descriptive statistics were generated and presented as mean (SD, range) for the demographic characteristics and for the amount of time spent on preparing, conducting and summarizing the supervision session. For the forced-choice questions and the visual analogue scales, frequencies and percentages (SD, range) were reported.

For Study II, III and IV, the inter-rater agreement of the MITI coding was assessed with a two-way mixed model with absolute agreement, single measures, ICC. Descriptive statistics were generated and presented as frequency, mean, standard error, median, inter-quartile range or percent. The primary outcomes in all these studies were the seven MITI summary values, the five MITI Beginning proficiency and Competency thresholds, the WAI-S summary score and its three subscales, and finally the additional question regarding discomfort and distress. Study II had three time points: Pre-training, post-training, and the six-month follow-up. Study III had only two: Post-training as the new baseline since the participants had there first supervision session after that recording, and the six-month follow-up. Study IV had had three time points: Baseline, the six-month assessment and the twelve-month follow-up.

To test the effectiveness of the different types of supervision and the supervisees' experience of the two different types of individual telephone supervision, a generalized linear mixed model (GLMM) was conducted for all theses studies. GLMM can handle missing data efficiently, as none of the participants excludes due to missing data at some point. The GLMM also provides flexibility as the choice of the covariance structures and links can be adapted to the nature of the data. All outcomes were examined independently for main effects (i.e., group, and time), and interaction (group X time). Using QQ-plots and other descriptive statistics, the distribution that most accurately represented data was chosen. The identity or the gamma link was used for the different covariance structures. Beyond nesting, repeated measures within individuals, and random intercept for individuals, other adaptations such as random slope or nesting within the councils did not provide a better fit for data. The Bonferroni correction was applied for multiple comparisons within each GLMM analysis, and the magnitude of the intervention effect was determined with Cohen's d. Chi-square analyses were employed to examine the differences of the proportion of participants who met the MITI Beginning proficiency and Competency thresholds in all three studies.

4 RESULTS

4.1 STUDY I

Five primary themes were identified in the analysis: (1) Structure of the sessions; (2) Educational methods and techniques; (3) Assessment of learning needs; (4) Supervision in an MI-consistent manner; and (5) Difficulties and opportunities. The results from the interviews in the first study also showed that many of the reported MI supervision behaviors were similar among the supervisors during supervision (e.g., all ten supervisors reported to verbally reinforce MI consistent behavior and to adjust behaviors that could be improved), but also that a number of behaviors varied (e.g., only five reported using role plays and only eight assigned homework to the supervisees). Moreover, the results showed that none of the ten supervisors felt that the MITI was relatively important for supervisees to learn MI, and five of the supervisors felt that MITI-based feedback might have had a negative effect on supervisees' learning.

4.2 STUDY II

In Study II, recruiting participants proved difficult, which may have led to a biased sample of practitioners highly motivated to learn the method. The study also had an attrition rate at 27.6 percent, and some participants did not always conduct the monthly recordings and supervision sessions according to the original study schedule. The most common stated reason for all these alterations was lack of time.

The analyses showed that, regardless of their slightly different form and content, and despite some differences in proficiency level at the pre-training assessment, all the five workshop trainings increased the participants' skills in MI to virtually the same level: Mean differences between the pre- and post-workshop assessment ranged from -0.06 to 3.49 across the proficiency measures, and the GLMM-analysis showed significant time effects for five of the seven MITI proficiency measures: Empathy ($F_{(1, 309)} = 67.31$, $p < .001$, $d = 1.46$), MI spirit ($F_{(1, 309)} = 78.28$, $p < .001$, $d = 1.58$), MI non-adherent behaviors ($F_{(1, 309)} = 119.26$, $p < .001$, $d = 1.95$), Reflection to question ratio ($F_{(1, 309)} = 42.96$, $p < .001$, $d = 1.19$), and Percent open questions ($F_{(1, 309)} = 13.71$, $p < .001$, $d = 0.66$).

Additionally, consistent with previous research, the additional supervision group showed larger gains in proficiency with significantly higher means for six of the seven MITI proficiency measures at the six-month follow-up. The GLMM-analyses also showed significant interactions effects for three of these seven MITI proficiency measures: MI spirit ($F_{(1, 266)} = 7.12$, $p < .01$, $d = 0.48$), Percent open questions ($F_{(1, 266)} = 5.06$, $p < .05$, $d = 0.40$), and Percent complex reflections ($F_{(1, 264)} = 14.77$, $p < .001$, $d = 0.68$). However, analyses showed generally maintained levels of skills for all participants at the follow-up assessment, and the majority of participants did not attain Beginning proficiency levels at either post-workshop or follow-up.

4.3 STUDY III

In Study III, the GLMM analyses of the seven MITI proficiency measures showed significant time and group interactions for two of the seven skill measures: MI non-adherent behaviors ($F_{(1,}$

152) $= 12.34$, $p = .001$, $d = 0.71$) and for Percent complex reflections ($F (1, 148) = 4.29$, $p = .040$, $d = 0.42$). The group that received systematic feedback based on only the behavioral component of MITI performed better at follow-up with regard to both of those proficiency measures, and thus had a better development over time. The group that received systematic feedback based on only the behavioral component of MITI also had a higher proportion of participants reaching Beginning proficiency levels on four of the five indicators at follow-up, but after the Bonferroni correction was applied, none of these differences remained significant.

Additionally, the analysis showed no between groups differences in the supervisor-supervisee working alliance, but the group that received systematic feedback based on the entire MITI protocol had a somewhat higher discomfort/distress score at baseline, and a somewhat lower score at follow-up, indicating a more apparent decrease of discomfort or feeling of distress over time for that group. However, at follow-up, the discomfort/distress scores had decreased in both groups from already low scores at baseline, indicating that the objective feedback was not difficult to handle for any of the supervisees

4.4 STUDY IV

In Study IV, 54 (40.3%) of the 134 participants chose not to complete their participation. However, independent t-tests showed no significant differences in primary outcome variables at baseline between those who did or did not complete the study.

Many participants met the benchmarks for Beginning proficiency already at baseline, indicating a successful implementation. Still, the results varied widely between the thresholds. Moreover, the regular group supervision and the supervision based on objective feedback were equally effective: The group means had increased from baseline to the six-month assessment for all proficiency measures except one in the regular group supervision group, and two in the individual telephone supervision group, and the GLMM-analyses showed no significant interaction or group effects. At the twelve-month follow-up, both groups' group means had then decreased again from the six-month assessment for all proficiency measures except for two.

After adjustments for baseline differences, the GLMM analyses of the two individual telephone supervision groups showed significant time and group interactions for one of the seven skill measures: Adherent behaviors ($F (1, 100) = 11.26$, $p < .001$, $d = 0.80$). The group that received systematic feedback based on only the behavioral component of MITI performed better at follow-up with regard to this single proficiency measures.

The results also indicated that the objective feedback was not difficult to handle for any of the supervisees in these two groups: The GLMM analyses of the working alliance and the discomfort/distress question showed no significant interaction or group effects, and also in this study, at follow-up, the discomfort/distress scores had decreased in both groups from already low scores at baseline. Moreover, just like Study II, Study IV had a high attrition rate (40.3%) and complications with participants not following the original study plan.

5 DISCUSSION

The overall objective of the present thesis was to evaluate various aspects of training of practitioners as part of the D&I of MI. In the first study, ten MI supervisors from a randomized controlled prevention study responded to one-to-one semi-structured telephone interviews about their supervision behaviors. The aim of the second study was to evaluate to what extent practitioners from a variety of publicly financed healthcare facilities, all with different strategies for the D&I, acquired skills from regular community-based MI workshop trainings, as opposed to trainings followed by supervision with form and content that previous studies have shown efficient. The third study evaluated the efficacy of two different ways to provide objective MI feedback during supervision, and the last study assessed the MI skills of practitioners with previous MI training in the context of an organization with a comprehensive MI implementation (SiS). Study IV also evaluated the effectiveness of the regular SiS subsequent group supervision, as opposed to trainings followed by supervision with form and content that previous studies have shown efficient, and replicated the third study on the efficacy of two different ways to provide objective MI feedback during supervision.

In the first study, many of the reported supervisory behaviors were in line with both the study instructions and with recommendations from the supervision literature (American Psychological Association, 2015; Carol A. Falender & Shafranske, 2007; Martino et al., 2008). However, the supervisors' drift away from key practices might have had a negative effect on supervisees' learning, which may have affected both the supervisees' skill acquisition and, indirectly, the intervention effects of the trial. The results from Study II and IV show that practitioners seems to learn a lot from both shorter community-based MI workshop trainings (Study II) and from MI trainings in the context of organizations with comprehensive MI implementation (Study IV). Yet, far from every practitioner reached the recommended levels of competence, which however needs to be explored further: The MITI thresholds for Beginning proficiency and Competency are only provisional expert-based cut-points along a continuum of skills, and the exact levels of MI proficiency sufficient for making a difference in client outcomes is unclear (Miller & Rollnick, 2014). The question of empirically derived fidelity standards also affects the recommendation of ongoing training until proficiency is achieved, instead of relying on fixed training doses (Beidas & Kendall, 2010; Hall et al., 2015; Martino et al., 2011; Miller & Moyers, 2017; Miller & Rollnick, 2014; Rakovshik & McManus, 2010).

Repeated studies have found that subsequent supervision should be included in workshop trainings for sustained changes in practitioners' actual behavior (Beidas et al., 2012; Beidas & Kendall, 2010; Edmunds et al., 2013; Herschell et al., 2010; Rakovshik & McManus, 2010; A. D. Roth et al., 2010; Webster-Stratton et al., 2014); a conclusion confirmed by the results in both Study II and IV. However, it is still unclear exactly how efficient supervision should be designed and provided (Reiser & Milne, 2012). The findings from Study I also suggest that for appropriate supervision activities to be conducted in each supervision session, clear supervision principles that specify the content and procedure of the supervision might be required, as well as some kind of adherence monitoring of the supervision sessions.

A lot is known about how people learn and how instructional designs can be modified to maximize learning, retention, and transfer of knowledge and skills (Benassi et al., 2014). These findings from the science of learning have clear implications for health education, but are often desolated from education in practice (Gooding et al., 2017). However, efforts to translate evidence-based approaches to learning to the health professions have emerged in recent years (Benassi et al., 2014; Gooding et al., 2017; Pashler et al., 2007). A training format with subsequent supervision is, in itself, spaced practice – a strategy with proven effectiveness for acquiring new learning and for enhancing long-term retention (Rohed & Paschler 2010). Additionally, recordings of practice samples requires the learner to recall information previously acquired – one of the cognitive strategies for learning with the strongest evidence base known as retrieval practice, or test-enhanced learning (Gooding et al., 2017). Recurring opportunities for supervisees to assess their knowledge with fidelity tools such as the MITI also provide them with frequent feedback on acquired skills, information that can give them information on what they still need to learn and assist them to engage more efficiently in further training (Cutting & Saks, 2012). Furthermore, supervision including feedback based on monitoring of sessions provides supervisees with opportunities for transfer of learning – the ability to apply acquired skills in a new and different context (Gooding et al., 2017; Young et al., 2014). However, it is still far from clear how to best use multifaceted feedback tools to promote learning during supervision.

Consistent with previous conclusions that supervisees can handle both positive and corrective performance feedback (Ellis, 2010; Ladany et al., 2013), both Study III and IV indicated that the objective feedback did not provoke supervisee discomfort/distress or adversely affect the supervisory relationship. However, although a restricted number of simultaneously presented variables might facilitate learning during supervision due to the limited capacity of the working memory (van Merrienboer & Sweller, 2010), the observed differences between the two groups MI skill acquisition were small in both Study III and IV (i.e., the groups receiving feedback on a restricted number of variables at each session performed better on only two of the seven skill measures in Study III, and one of the seven skill measures in Study IV). It is also not clear what really generated these differences. According to CLT (van Merrienboer & Sweller, 2010), a restriction of the number of variables at each session when providing objective feedback from multifaceted feedback tools can possible be a more effective way to promote learning, at least for novice supervisees. Part-task or sequenced practice also allows for repetition and for the supervisee to reflect and practice in between sessions, activities all of which have proven to be important for the acquisition of long-term knowledge and skills (Kerfoot et al., 2007; Rakovshik & McManus, 2010; Young et al., 2014), and training provided in doses could thereby maybe be a more effective form of learning.

Additionally, research on interleaving has shown that if a restricted number of key variables are presented at each session, it is important to get a mixed presentation of the supervision content in an alternating fashion to promote integration of learning and long-term retention (Rohrer & Pashler, 2010). Moreover, the expertise reversal effect (Kalyuga, 2007) has demonstrated the importance of adjusting instructional methods and procedures as learners acquire more expertise in a specific domain: While novice learners dealing with many new elements of information may easily overload their working memory, more expert learners can rely on their available long-term memory for handling situations and tasks within their area of

expertise (Kalyuga, 2007). Cumulative adjustments in accordance with the development of supervisee learning might thereby be an even more sufficient strategy than just restricting the number of variables presented at each supervision session, when using multifaceted feedback tools in supervision.

An additional way to use fidelity tools, such as the MITI, for enhanced learning during supervision is to promote metacognition, or thinking about thinking. Metacognition is critical for learners to manage their own learning in any domain, and earlier research has shown that it becomes increasingly important as students advance (Benassi et al., 2014). However, repeated studies have also shown that learners' ability to assess their own accomplishments is limited (Beidas & Kendall, 2010; Benassi et al., 2014; Hogue, Dauber, Lichvar, Bobek, & Henderson, 2015; Martino, Ball, Nich, Frankforter, & Carroll, 2009; Rohrer & Pashler, 2010; Wain et al., 2015), but when they are repeatedly tested, their estimates become more accurate (Benassi et al., 2014). Giving the supervisees opportunities to self-assess their abilities based on fidelity tools might be another way to further enhance this ability as part of regular supervisee practice.

5.1 LIMITATIONS

The included studies contain a number of limitations. In Study I, the supervisors' behaviors were self-reported and not based on observations or objective measures of performance with a standardized instrument with known psychometric properties, which may raise questions about both the reliability and validity of the results. However, a significant number of open-ended questions and a mixed method approach were deemed necessary because of the exploratory nature of the study. In future supervision studies the supervisory sessions would ideally be recorded in order to provide objective data.

In Study II and III, the sample of self-selected participants may not well represent the larger population of county council practitioners, and in Study IV the population of practitioners working with compulsory care for people with substance use disorders (SUD), which limits the generalizability of findings in all these three studies. Moreover, when studying the effectiveness of the county council workshop trainings in Study II, the absence of a comparison group and the fact that the participants were not randomized to the five groups presents a number of threats to the study's internal validity. Study III and IV also lacked control groups (i.e., a supervision group without objective feedback in Study III and IV, and also a group without supervision in Study IV).

The use of standardized patients, as in Study II and III, does not provide adequate information about how MI is employed in actual clinical practice; Decker and colleagues (2013) found that clinicians were significantly more MI adherent and used more advanced MI strategies in role-played sessions than in real client sessions, and demonstrated poor rating correspondence between the two assessment approaches. Yet, standardized patients allow for clients' characteristics to be kept constant between both participants and assessment points, indicating

that, if not used interchangeably with real client sessions, role-played sessions can provide useful information about therapists' MI performance (Decker et al., 2013; Imel et al., 2014).

In Study IV, all recordings could not be conducted according to the study plan (i.e., together with clients) and there was a significant difference ($p = 0.011$) regarding choice of client or colleague as recording partner between the participants in the two supervision groups at the six-month assessment. However, choice of client or colleague as a moderator in the later analyses showed no significant interaction effects between the two groups over time for any of the outcome measures.

The follow-up assessment took place six months after the post-workshop assessment in Study II, and twelve months after the baseline recording in Study IV. Longer follow-up periods would have given a better picture of how participants' MI skills are sustained over time. There was also only one recording per assessment point in both these studies and in Study III. Since MI performance often varies significantly within therapists (Dunn et al., 2016; Imel et al., 2014), repeated measures at each assessment point would have assured a more accurate estimates of the participants' MI integrity. Furthermore, the coders in all these three studies were not blind to the participants' group allocation during coding, and the same sample of coders rated sessions and performed supervision with the participants, which may have affected the reliability of the coding (Moyers et al., 2016).

Another limitation is that, even if the supervisors had supervisory meetings every quarter that included self-selected samples, and all telephone supervision sessions were recorded during a three-month period in Study II and III, the supervisors were never assessed for feedback fidelity in any of the included studies. It is thus not clear if the supervisors could refrain from revealing the MITI BC group's global scores in Study III and IV, or how the MI-consistent supervision style affected the impact of the different types of feedback in Study II, III and IV.

Many of the limitations in Study II, III and IV derive from the complexity of conducting studies in naturalistic settings. Within the scope of all these studies, it was not possible to also evaluate the impact of organizational-level variables. However, the recruitment difficulties in all these trials had consequences for the ideal sample sizes. All three trials also had problem with participants dropping out. The most common stated reason for not participating or completing the studies was time constraints. Participants who completed the studies also frequently reported time pressures and, in some instances, insufficient organizational support in Study II and III. These contextual challenges for learning transfer should be addressed in future dissemination studies.

Despite these limitations, the included studies contributes to both the knowledge of MI training and subsequent supervision, and on how to provide objective feedback from multifaceted feedback tools, all in naturalistic settings with relatively large samples of practitioners, and can thereby provide some direction and considerations for future MI D&I studies.

28

5.2 CONCLUSION

This thesis addresses the real-life implications of D&I of EBT. In accordance with previous research, both Study II and IV showed that workshop training can increase participants' MI skills, and that subsequent supervision can further enhance acquired skills. Additionally, both studies indicate that these results also apply to naturalistic settings. However, the high variation in competence at all assessment points, together with the low interest in the possibility of additional supervision in both these studies, are troublesome. In addition, neither the workshop trainings, nor the costly additional sessions of individual telephone supervision, or the comprehensive MI-implementation within SIS, were sufficient for many of the participants to reach beginning proficiency levels. Although the levels of MI proficiency sufficient for making a difference in client outcomes is unclear (Miller & Rollnick, 2014), this raises questions regarding both the most efficient form of training for practitioners to attain and sustain adequate practice standards, and how to create an interest among practitioners to participate in such training.

Even though restricting the number of variables when providing objective feedback might promote learning during supervision, results from Study III and IV could not conclude that MI supervision based on fewer MITI variables promotes supervisees' skill acquisition. However, the results from both these studies indicate that objective feedback does not seem to provoke significant supervisee anxiety, or negatively affect the supervisory relationship. Nonetheless, variation in the use of supervision components that previous research has proposed to be potentially influential to the process and outcome may lead to less efficient supervision. Findings from Study I suggest that appropriate supervision activities conducted in each supervision session require clear supervision principles that specify the content and procedure of the supervision, as well as regular adherence monitoring of the supervision sessions.

Given the high costs and resources associated with training practitioners as part of the D&I, and to ensure that patients receive intended and proper care, health care and community organizers should take available research knowledge into account when designing training programs for practitioners. To date, findings from health care science suggest that the most effective training strategies for sufficient and sustained skill acquisition among practitioners include subsequent supervision. However, questions remain regarding exactly how supervision should be designed and provided in order for it to have lasting and meaningful impact on the supervisee. One such question is how to best deliver feedback from multifaceted feedback tools. The more general science of learning have clear implications also for training practitioners, and can provide information regarding learning and how instructional designs can be modified to maximize retention and transfer of knowledge and skills. Organizers should take also this knowledge into account when designing training for practitioners. Additionally, future research needs to determine the effects of findings from the science of learning in a variety of educational healthcare contexts; additional research necessary for answering the many remaining questions regarding which training methods are most efficient in the context of training practitioners. Another question that still remains is the thresholds for competent practice. For the recommendations on ongoing training until proficiency is achieved to be meaningful, levels of proficiency sufficient for making difference in client outcomes must be clear. There are obvious both ethical and economic

values in ensuring that interventions are delivered according to given standards. Effective and affordable methods for assessing provider skills, and ways to match type and amount of training to distinct participant training needs, are all additional important aspects of future studies.

6 ACKNOWLEDGEMENTS

7 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*(1), 2-14. doi:10.1080/15401383.2011.560067

American Psychological Association. (2014). Guidelines for Clinical Supervision in Health Service Psychology.   Retrieved from http://apa.org/about/policy/guidelines-supervision.pdf

American Psychological Association. (2015). Guidelines for clinical supervision in health service psychology. *American Psychologist, 70*(1), 33-46. doi:10.1037/a0038112

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*(9), 1786-1795. doi:10.1016/j.childyouth.2012.05.012

Bearman, S. K., Schneiderman, R. L., & Zoloth, E. (2016). Building an Evidence Base for Effective Supervision Practices: An Analogue Experiment of Supervision to Increase EBT Fidelity. *Administration and Policy in Mental Health*. doi:10.1007/s10488-016-0723-8

Beidas, R. S., Edmunds, J. M., Marcus, S. C., & Kendall, P. C. (2012). Training and consultation to promote implementation of an empirically supported treatment: a randomized trial. *Psychiatric Services, 63*(7), 660-665. doi:10.1176/appi.ps.201100401

Beidas, R. S., & Kendall, P. C. (2010). Training Therapists in Evidence-Based Practice: A Critical Review of Studies From a Systems-Contextual Perspective. *Clinical Psychology, 17*(1), 1-30. doi:10.1111/j.1468-2850.2009.01187.x

Benassi, V. A., Overson, C. E., & Hakala, C. M. (2014). *Applying science of learning in education: Infusing psychological science into the curriculum.*

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

Bogo, M., & McKnight, K. (2006). Clinical Supervision in Social Work. *The Clinical Supervisor, 24*(1-2), 49-67. doi:10.1300/J001v24n01_04

Borders, L. D., DeKruyf, L., Fernando, D. M., Glosoff, H. L., Hays, D. G., Page, B., & Welfare, L. E. (2011). Best practices in clinical supervision. Retrieved from http://www.acesonline.net/resources/best-practices-clinical-supervision

Bosse, H. M., Mohr, J., Buss, B., Krautter, M., Weyrich, P., Herzog, W., . . . Nikendei, C. (2015). The benefit of repetitive skills training and frequency of expert feedback in the early acquisition of procedural skills. *BMC Medical Education, 15*, 22. doi:10.1186/s12909-015-0286-5

Chur-Hansen, A., & McLean, S. (2006). On being a supervisor: the importance of feedback and how to give it. *Australas Psychiatry, 14*(1), 67-71. doi:10.1080/j.1440-1665.2006.02248.x

Cicchetti, & Sparrow. (1981). Developing criteria for establishing interrater reliability of specific items: Applications to assessment of adaptive behavior. *American Journal of Mental Deficiency, 86 (2)*, 127–137.

Cicchetti, D. (1994). Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology. *Psychological Assessment, 6*(4), 284-290.

Clarke, P. B., & Giordano, A. L. (2013). The Motivational Supervisor: Motivational Interviewing as a Clinical Supervision Approach. *The Clinical Supervisor, 32*(2), 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*(2), 158-163. doi:10.1037/a0038712

Cutting, M. F., & Saks, N. S. (2012). Twelve tips for utilizing principles of learning to support medical education. *Medical Teacher, 34*(1), 20-24. doi:10.3109/0142159X.2011.558143

D. Friedberg, R., Mahr, S., & Mahr, F. (2010). Training Psychiatrists in Cognitive Behavioral Psychotherapy: Current Status and Horizons. *Current Psychiatry Reviews, 6*(3), 159-170. doi:10.2174/157340010791792563

Damschroder, L. J., Aron, D. C., Keith, R. E., Kirsh, S. R., Alexander, J. A., & Lowery, J. C. (2009). Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci, 4*, 50. doi:10.1186/1748-5908-4-50

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*(2), 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*(1), 306-312. doi:10.1037/a0030815

Doring, N., Hansson, L. M., Andersson, E. S., Bohman, B., Westin, M., Magnusson, M., . . . Rasmussen, F. (2014). Primary prevention of childhood obesity through counselling sessions at Swedish child health centres: design, methods and baseline sample characteristics of the PRIMROSE cluster-randomised trial. *BMC Public Health, 14*, 335. doi:10.1186/1471-2458-14-335

Dorsey, S., Kerns, S. E. U., Lucid, L., Pullmann, M. D., Harrison, J. P., Berliner, L., . . . Deblinger, E. (2018). Objective coding of content and techniques in workplace-based supervision of an EBT in public mental health. *Implement Sci, 13*(1), 19. doi:10.1186/s13012-017-0708-3

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*(2), 152-165. doi:10.1111/cpsp.12031

Elliott, R., Bohart, A. C., Watson, J. C., & Greenberg, L. S. (2011). Empathy. *Psychotherapy (Chic), 48*(1), 43-49. doi:10.1037/a0022187

Ellis, M. V. (2010). Bridging the Science and Practice of Clinical Supervision: Some Discoveries, Some Misconceptions. *The Clinical Supervisor, 29*(1), 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*(4), 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*(3), 232-240. doi:10.1037/0735-7028.38.3.232

Falender, C. A., & Shafranske, E. P. (2014). Clinical supervision: the state of the art. *Journal of Clinical Psychology, 70*(11), 1030-1041. doi:10.1002/jclp.22124

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). Implementation Research: A Synthesis of the Literature. .

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 3_1 SV.pdf

Forsberg, L., Kallmen, H., Hermansson, U., Berman, A. H., & Helgason, A. R. (2007). Coding counsellor behaviour in motivational interviewing sessions: inter-rater reliability for the Swedish Motivational Interviewing Treatment Integrity Code (MITI). *Cognitive Behaviour Therapy, 36*(3), 162-169. doi:10.1080/16506070701339887

Friedberg, R., Gorman, A., & Beidel, D. (2009). Training Psychologists for Cognitive-Behavioral Therapy in the Raw World - A Rubric for Supervisors. *Behavior Modification, 33*(1), 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*(1), 26-31. doi:10.1080/0142159X.2016.1231913

Grimshaw, J. M., Eccles, M. P., Lavis, J. N., Hill, S. J., & Squires, J. E. (2012). Knowledge translation of research findings. *Implement Sci, 7*, 50. doi:10.1186/1748-5908-7-50

Hall, K., Staiger, P. K., Simpson, A., Best, D., & Lubman, D. I. (2015). After 30 years of dissemination, have we achieved sustained practice change in motivational interviewing? *Addiction*. doi:10.1111/add.13014

Herschell, A. D., Kolko, D. J., Baumann, B. L., & Davis, A. C. (2010). The role of therapist training in the implementation of psychosocial treatments: a review and critique with recommendations. *Clinical Psychology Review, 30*(4), 448-466. doi:10.1016/j.cpr.2010.02.005

Hogue, A., Dauber, S., Lichvar, E., Bobek, M., & Henderson, C. E. (2015). Validity of therapist self-report ratings of fidelity to evidence-based practices for adolescent behavior problems: correspondence between therapists and observers. *Administration and Policy in Mental Health, 42*(2), 229-243. doi:10.1007/s10488-014-0548-2

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research, 15*(9), 1277-1288. doi:10.1177/1049732305276687

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*(3), 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 Syst Rev*(6), CD000259. doi:10.1002/14651858.CD000259.pub3

Johnson, W. B., & Kaslow, N. J. (Eds.). (2014). *The Oxford Handbook of Education and Training in Professional Psychology*. New York, NY: Oxford University Press.

Kalyuga, S. (2007). Expertise Reversal Effect and Its Implications for Learner-Tailored Instruction. *Educational Psychology Review, 19*(4), 509-539. doi:10.1007/s10648-007-9054-3

Kazdin, A. E. (2008). Evidence-based treatment and practice: new opportunities to bridge clinical research and practice, enhance the knowledge base, and improve patient care. *American Psychologist, 63*(3), 146-159. doi:10.1037/0003-066X.63.3.146

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*(4), 1481-1487. doi:10.1016/j.juro.2006.11.074

Kirk, M. A., Kelley, C., Yankey, N., Birken, S. A., Abadie, B., & Damschroder, L. (2016). A systematic review of the use of the Consolidated Framework for Implementation Research. *Implement Sci, 11*, 72. doi:10.1186/s13012-016-0437-z

Ladany, N., Mori, Y., & Mehr, K. E. (2013). Effective and Ineffective Supervision. *The Counseling Psychologist, 41*(1), 28-47. doi:10.1177/0011000012442648

Liese, B. S., & Beck, J. S. (1997). *Cognitive therapy supervision. In C. E. Watkins (Ed.) Handbook of psychotherapy supervision* Hoboken, NJ, US: John Wiley & Sons Inc.

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*(2), 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*(1), 101-109. doi:10.1016/j.jsat.2008.05.005

Magill, M., Apodaca, T. R., Borsari, B., Gaume, J., Hoadley, A., Gordon, R. E. F., . . . Moyers, T. (2018). A meta-analysis of motivational interviewing process: Technical, relational, and conditional process models of change. *Journal of Consulting and Clinical Psychology, 86*(2), 140-157. doi:10.1037/ccp0000250

Martino, S., Ball, S., Nich, C., Frankforter, T. L., & Carroll, K. M. (2009). Correspondence of motivational enhancement treatment integrity ratings among therapists, supervisors, and observers. *Psychother Res, 19*(2), 181-193. doi:10.1080/10503300802688460

Martino, S., Canning-Ball, M., Carroll, K. M., & Rounsaville, B. J. (2011). A criterion-based stepwise approach for training counselors in motivational interviewing. *Journal of Substance Abuse Treatment, 40*(4), 357-365. doi:10.1016/j.jsat.2010.12.004

Martino, S., Gallon, S., Ball, S., & Carroll, K. M. (2008). A Step Forward in Teaching Addiction Counselors How to Supervise Motivational Interviewing Using a Clinical Trials Training Approach. *Journal of Teaching in the Addictions, 6*(2), 39-67. doi:10.1080/15332700802127946

Miller, & Moyers. (2017). Motivational interviewing and the clinical science of Carl Rogers. *Journal of Consulting and Clinical Psychology, 85*(8), 757-766. doi:10.1037/ccp0000179

Miller, & Rollnick. (2013). *Motivational interviewing: Helping people change*. New York: Guilford Press.

Miller, & Rollnick. (2014). The effectiveness and ineffectiveness of complex behavioral interventions: impact of treatment fidelity. *Contemporary Clinical Trials, 37*(2), 234-241. doi:10.1016/j.cct.2014.01.005

Miller, W. R., & Rose, G. S. (2009). Toward a theory of motivational interviewing. *American Psychologist, 64*(6), 527-537. doi:10.1037/a0016830

Milne, D., & Reiser, R. P. (2012). A Rationale for Evidence-Based Clinical Supervision. *Journal of Contemporary Psychotherapy, 42*(3), 139-149. doi:10.1007/s10879-011-9199-8

Moyers, Martin, Manuel, Hendrickson, & Miller. (2005). Assessing competence in the use of motivational interviewing. *Journal of Substance Abuse Treatment, 28*(1), 19-26. doi:10.1016/j.jsat.2004.11.001

Moyers, Martin, Manuel, Miller, & Ernst. (2010). Revised global scales: Motivational interviewing treatment integrity 3.1.1 (MITI 3.1.1) Retrieved from http://casaa.unm.edu/mimanuals.html

Moyers, Rowell, Manuel, Ernst, & Houck. (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

Moyers, T. B., & Miller, W. R. (2013). Is low therapist empathy toxic? *Psychology of Addictive Behaviors, 27*(3), 878-884. doi:10.1037/a0030274

Nadeem, E., Gleacher, A., & Beidas, R. S. (2013). Consultation as an implementation strategy for evidence-based practices across multiple contexts: unpacking the black box. *Administration and Policy in Mental Health, 40*(6), 439-450. doi:10.1007/s10488-013-0502-8

Newman, C. (2010). Competency in conducting cognitive-behavioral therapy: foundational, functional, and supervisory aspects. *Psychotherapy Theory, Research, Practice, Training, 47*(1), 12-19. doi:10.1037/a0018849

Pace, B. T., Dembe, A., Soma, C. S., Baldwin, S. A., Atkins, D. C., & Imel, Z. E. (2017). A multivariate meta-analysis of motivational interviewing process and outcome. *Psychology of Addictive Behaviors, 31*(5), 524-533. doi:10.1037/adb0000280

Padesky, C. A. (1996). *Developing cognitive therapist competency: Teaching and supervision models. In P. M. Salkovskis (Ed.), Frontiers of cognitive therapy* New York, NY, US: Guilford Press.

Parsons, M. B., Rollyson, J. H., & Reid, D. H. (2012). Evidence-based staff training: a guide for practitioners. *Behav Anal Pract, 5*(2), 2-11.

Pashler, H., Rohrer, D., Cepeda, N. J., & Carpenter, S. K. (2007). Enhancing learning and retarding forgetting: Choices and consequences. *Psychonomic Bulletin & Review, 14*(2), 187-193. doi:10.3758/bf03194050

Powell, B. J., McMillen, J. C., Proctor, E. K., Carpenter, C. R., Griffey, R. T., Bunger, A. C., . . . York, J. L. (2012). A compilation of strategies for implementing clinical innovations in health and mental health. *Medical Care Research and Review, 69*(2), 123-157. doi:10.1177/1077558711430690

Powell, B. J., Proctor, E. K., & Glass, J. E. (2014). A Systematic Review of Strategies for Implementing Empirically Supported Mental Health Interventions. *Res Soc Work Pract, 24*(2), 192-212. doi:10.1177/1049731513505778

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*(5), 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*(3), 191-199. doi:10.1037/ccp0000079

Reiser, R. P., & Milne, D. (2012). Supervising Cognitive-Behavioral Psychotherapy: Pressing Needs, Impressing Possibilities. *Journal of Contemporary Psychotherapy, 42*(3), 161-171. doi:10.1007/s10879-011-9200-6

Rohrer, D., & Pashler, H. (2010). Recent Research on Human Learning Challenges Conventional Instructional Strategies. *Educational Researcher, 39*(5), 406-412. doi:10.3102/0013189x10374770

Rosenbaum, M., & Ronen, T. (1998). Clinical supervision from the standpoint of cognitive-behavior therapy. *Psychotherapy, 35*(2), 220-230.

Roth, A., & Pilling, S. (2009). A competence framework for the supervision of psychological therapies.

Roth, A. D., Pilling, S., & Turner, J. (2010). Therapist training and supervision in clinical trials: implications for clinical practice. *Behavioural and Cognitive Psychotherapy, 38*(3), 291-302. doi:10.1017/S1352465810000068

Schmidt, R. A., & Bjork, R. A. (2012). New Conceptualizations of Practice: Common Principles in Three Paradigms Suggest New Concepts for Training. *Psychological Science, 3*(4), 207-218. doi:10.1111/j.1467-9280.1992.tb00029.x

Schwalbe, C. S., Oh, H. Y., & Zweben, A. (2014). Sustaining motivational interviewing: a meta-analysis of training studies. *Addiction, 109*(8), 1287-1294. doi:10.1111/add.12558

Sewell, K. (2017). Theoretically grounded, evidence-informed clinical supervision for the SNAP programs: A model in development. *The Clinical Supervisor, 36*(2), 340-359. doi:10.1080/07325223.2017.1352549

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 cognitive-behavioral therapy. *Journal of Consulting and Clinical Psychology, 73*(1), 106-115. doi:10.1037/0022-006X.73.1.106

Soderlund, L. L., Madson, M. B., Rubak, S., & Nilsen, P. (2011). A systematic review of motivational interviewing training for general health care practitioners. *Patient Education and Counseling, 84*(1), 16-26. doi:10.1016/j.pec.2010.06.025

Tabak, R. G., Khoong, E. C., Chambers, D. A., & Brownson, R. C. (2012). Bridging research and practice: models for dissemination and implementation research. *American Journal of Preventive Medicine, 43*(3), 337-350. doi:10.1016/j.amepre.2012.05.024

The Psychology Board of Australia. (2013). Guidelines for supervisors and supervisor training providers.

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. *Behav Anal Pract, 9*(4), 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*(1), 85-93. doi:10.1111/j.1365-2923.2009.03498.x

Wain, R. M., Kutner, B. A., Smith, J. L., Carpenter, K. M., Hu, M. C., Amrhein, P. C., & Nunes, E. V. (2015). Self-Report After Randomly Assigned Supervision Does not Predict Ability to Practice Motivational Interviewing. *Journal of Substance Abuse Treatment, 57*, 96-101. doi:10.1016/j.jsat.2015.04.006

Watkins, C. E. (2011). Does Psychotherapy Supervision Contribute to Patient Outcomes? Considering Thirty Years of Research. *The Clinical Supervisor, 30*(2), 235-256. doi:10.1080/07325223.2011.619417

Watkins, C. E. (2012a). Contemporary Visions of Psychotherapy Supervision: Sharing Perspective, Identifying Need, and Charting Possibility. *Journal of Contemporary Psychotherapy, 42*(3), 125-127. doi:10.1007/s10879-011-9203-3

Watkins, C. E. (2012b). Psychotherapy Supervision in the New Millennium: Competency-Based, Evidence-Based, Particularized, and Energized. *Journal of Contemporary Psychotherapy, 42*(3), 193-203. doi:10.1007/s10879-011-9202-4

Watkins, E. (2014). Clinical Supervision in the 21st Century: Revisiting Pressing Needs and Impressing Possibilities. *American Journal of Psychotherapy, 68*(2), 251-272.

Webster-Stratton, C. H., Reid, M. J., & Marsenich, L. (2014). Improving therapist fidelity during implementation of evidence-based practices: Incredible years program. *Psychiatric Services, 65*(6), 789-795. doi:10.1176/appi.ps.201200177

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*(1), 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*(5), 371-384. doi:10.3109/0142159X.2014.889290

8 APPENDIX

8.1   THE TELEPHONE SUPERVISION MANUAL

1. Code the recording and write your comments (just as a regular coding).

2. Use boxes below for notes to support the supervision session.

Order number: Session date and time: Telephone number:

Session keywords (e.g., client gender).

What did the supervisee do well?

Room for improvement (1-2 areas).

Session practice samples (a minimum of two role-plays). Note recording time code.

3. Email the protocol to the supervisee.
Ten minutes before the session, email the protocol to the supervisee without adding comments. Use the remaining time to prepare for the session.

> *Email template:*
> *Hi first name,*
>
> *We have a schedule telephone supervision session now at xx:xx am/pm. Attached is your coded protocol without comments. Please review the protocol and prepare any questions and/or comments.*
>
> *At xx:xx am/pm, I will call you at this number (xxx) xxx-xxxx.*
>
> *Kind regards, Your name*

4. The supervision session

General information:

Lengths for each part are approximate, but it is important that all parts are included. Primary focus should be on the role-plays (D).

During the session, consider two aspects: 1) What the supervisee needs to improve based on the coding; 2) What the supervisee thinks is difficult / wants to focus on.

If this is the first time that the supervisee receives MITI-based feedback, the supervisee needs an overview of the MITI protocol. Remember that some supervisees only get the behavior counts part of the protocol!

A) Introduction (about 5 minutes)

Introduce yourself and plan the session together with the supervisee (i.e., set the agenda). Ask specifically about session objectives or previous practice focus / homework.

B) What did the supervisee do well (about 5 minutes)

Start by eliciting supervisee's thoughts on what worked well, adding your own notes. Verbally reinforce progress / MI consistent utterances.

C) Room for improvement (about 5 minutes)

Start by eliciting supervisee's thoughts on difficulties / things that can be improved, adding your own notes. Strive for one to two practice areas. Joint problem solving / elicit suggestions for alternative behaviors from the supervisee.

D) Practice (about 10 minutes)

Use the practice samples from the session, and work back and forth with the supervisee (i.e., modeling of MI consistent behaviors *and* supervisee practice).

E) Homework (about 3 minutes)

Start by eliciting supervisee's suggestions, adding the session practice areas. Strive for no more than one or two areas of homework.

F) Summary and evaluation (2 min)

Ask the supervisee to summarize and evaluate the session (e.g., something in today's session that went well / that was useful / could be improved).

Before the session ends, inform the supervisee when and how the protocol can be downloaded from the website.

8.2 QUESTION 9 TO 32 OF THE INTERVIEW PROTOCOL USED IN STUDY I

9	Below you will find several possible starting points for the agenda setting. Please rate them in terms of how important they are to you when structuring your sessions by ranking them in descending order. You may also add starting points important to you. The MITI coding Requests from the supervisee The audio recorded practice sample Feedback from previous sessions ……………………………………….
10	Please outline the way you usually structure your supervision sessions. List the main features in the order you utilize them, and report the approximate time required for each one.
11	What component(s) of supervision do you think is/are most important for the supervisee's acquisition of MI skills? If more than one component, please rank them in descending order.
12	Do you think any component(s) of supervision might have a negative effect on the supervisee's MI's skill acquisition? If more than one aspect, please rank them in descending order.
13	To what extent do you usually practice supervision in a style consistent with MI? Please choose a number between 0 (not at all) and 100 (to a very large extent).
14	Which MI principle(s) and/or strategy(ies) do you typically use in your supervision? If more than one principle and/or strategy, please rank them in descending order.
15	What is your most crucial tool for identifying the strengths in the supervisee's MI practice? Please check off one alternative below or describe a tool of your own. The MITI coding The audio recorded practice sample The supervisee's judgment of their practice sample Other, please specify ………………………………………………………….

16	What is your most crucial tool for identifying the weaknesses in the supervisee's MI practice? Please check off one alternative below or describe a tool of your own.
	The MITI coding
	The audio recorded practice sample
	The supervisee's judgment of their practice sample
	Other, please specify
	………………………………………………………….
17	Please rate your balance in focus on the supervisee's strengths and difficulties during sessions. Estimate the balance in per cent (with the total sum 100).
	………………% Strengths ……………..% Difficulties
18	What rationale do you provide to the supervisee to help them use the ratings in the MITI protocol in a constructive way?
19	How do you use the MITI as part of your supervision practice?
20	When and how do you provide the MITI protocol outcome to the supervisee?
21	When listening to the recorded practice sample, to what extent do you usually focus on what the supervisee says in relation to what the client says? Estimate the balance in per cent (with the total sum 100).
	………………% Focus on supervisee ……………..% Focus clients
22	Do you think 30 minutes is enough time to prepare ahead of the supervision sessions? Please check off one of the alternatives. YES NO
23	Do you find 30 minutes is enough time for the supervision sessions? Please check off one of the alternatives. YES NO
24	Do you find 30 minutes is enough time to compile the supervision sessions? Please check off one of the alternatives. YES NO
25	How many minutes of the recorded practice samples do you usually listen to?

26	How many sequences of the recorded practice samples do you usually listen to?
27	What part of the recorded practice samples do you usually listen to?
28	Please explain the reasons why supervision sessions are not always conducted according to the time instructions. Reasons related to me as supervisor Reasons related to the supervisee
29	How do you think telephone supervision compares to face-to-face supervision? Please choose a number between 0 (not at all) and 100 (to a very large extent).
30	Do you think MI is consistent with a manual-based intervention, such as PRIMROSE? Please check off one of the alternatives. Not consistent Partially consistent Fully consistent
31	Are the monthly telephone supervisory meetings helpful for your supervision practice? Please rate their helpfulness from 0 (not at all) and 100 (to a very large extent).
32	In what way do you find the monthly telephone supervisory meetings helpful to your supervision practice? Please provide a brief answer. If more than one way, please rank them in descending order.
	Additional comments and suggestion: ……………………………………………………………………………………………………….