Measuring MBT – A marriage of the common and specific psychotherapy factors

Espen Jan Folmo

UNIVERSITY OF OSLO

January 2021

Table of contents

1 Summary

Background: Borderline personality disorder (BPD) is a severe condition traditionally considered difficult to treat and is characterized by both self and relational pathology. Attachment and mentalization theory considers the poorly developed capacity for stable, close relationships and (epistemic) trust as essential aspects of BPD. In terms of the therapeutic relationship, the development and maintenance of the alliance constitutes a central issue in therapy and a curing mechanism for BPD patients. However, it is well known that the therapeutic alliance can be extremely complicated in the treatment of BPD. There are also indications that patients with poor alliance are more likely to drop out early in therapy. BPD patients typically have a dropout rate of around 29%, but the real number may be higher due to bias towards the publication of studies with high completion rates. Importantly, different evidence-based treatments for BPD seem to work through, and create different, alliances. Therefore, the therapeutic alliance seems a crucial factor both for keeping patients in therapy and for effective treatment. The impact of alliance for patients with personality disorders (PDs) has been shown to be six times higher than for other patient groups. Despite the challenges involved in psychotherapy with individuals with BPD, there exists little literature on the therapeutic alliance in this population. Therefore, as the alliance, the quintessential common factor (CF) between treatments, has been poorly investigated in BPD treatment, this thesis examined the role of the working alliance and the CFs in an evidence-based treatment for BPD. An overall question in the thesis is how psychotherapy research can build a bridge between the so-called CFs and specific factors (i.e., the current schism). As reliable fidelity measures are necessary to judge a specific treatment technique as being superior to another, this thesis also aimed to establish reliable integrity measures for mentalization-based treatment (MBT).

Objectives: Paper I examined the reliability of a measure of treatment fidelity for the group component of MBT. Paper II investigated whether differences in rated MBT quality can be investigated through the lens of the CFs; for example, can therapeutic alliance account for some of the differences between high and poor ratings of therapists' MBT fidelity? Paper III studied how aspects of therapeutic alliance (goals, tasks, and bonds) have developed over time in MBT for patients with BPD with different outcomes.

Methods: Paper I applied generalizability theory (G-theory) in a reliability study (G-study) where five raters rated eight MBT group (MBT-G) and eight psychodynamic group (PDG) sessions according to the newly developed adherence and quality scale for MBT-G (MBT-G-AQS). Paper II applied purposeful sampling to a pool of 108 rated sessions in the Quality Lab for Psychotherapy at Oslo University Hospital. The four selected sessions were subject to interpretative phenomenological analysis. Paper III applied linear mixed models to investigate the longitudinal development of alliance for 155 BPD patients in MBT. Psychosocial functioning measured using the Global Assessment of Functioning (GAF) scale indicated clinical outcomes. Subscales (goals, tasks, and bonds) of the working alliance as measured by the Working Alliance Inventory-Short Revised (WAI-SR) were the dependent variables in an analysis of alliance development in MBT therapies with good (end-GAF < 60) and poor outcomes (end-GAF ≥ 60).

Results: The results in Paper I showed high reliability for both adherence and quality (competence). The mean absolute G-coefficient for adherence was .86 (range .63–.97) and for quality was .88 (range .64–.96) with five raters. The reliability for overall adherence (.97) and quality (.96) ratings (five raters) were both excellent. The nine group-specific items displayed high reliability for both adherence (range .83–.95) and quality (range .78–.96). With one rater, the reliability was also high for overall MBT-G adherence (.86) and quality (.83). However, the results indicated low reliability for items connected to psychic equivalence and pretend mode, especially with few raters. Paper II identified four themes that seemed to characterize therapy processes with different ratings of MBT quality (competence): 1) alliance, 2) strategic competence, 3) quality, and 4) "battles of the comfort zone". Therapeutic alliance seemed to be fostered by battles of the comfort zone, quality, and strategy. Given an existing, adequate alliance between patient and therapist, the alliance seemed to become further nurtured when the interventions targeted the patients' maladaptive patterns. Highly rated therapists intervened according to an overarching strategy and challenged the patients' comfort zone. They also manifested a steadfast focus on the agreed therapeutic project (tasks and goals). The bond part of the alliance appeared as an asset in this "battling" process. Poorly rated therapists abandoned the therapeutic project and seemed overwhelmed by countertransference reactions. The therapeutic strategy seemed random and the sessions had low levels of challenging maladaptive patterns. Paper III showed that MBT treatments with good outcomes were characterized by positive development in the working alliance. Differences between subgroups with good and poorer outcomes were most prominent for the tasks subscale. Initial

ratings of goals, bonds, and tasks did not differ by subgroup; levels were within a satisfactory range, but change over time was significantly different by subgroup. Comorbid paranoid PD was more frequent in the subgroup with poor outcomes and associated with poorer alliance development in this subgroup. Mood disorder was associated with significantly lower initial alliance levels but not with change in the working alliance subscales.

Conclusions: The overall MBT-G adherence and quality can be rated by one rater. Two of the core components of MBT theory, psychic equivalence and pretend mode, had low reliability. Paper II found that MBT-I may foster a strong therapeutic alliance and that CFs, such as alliance, believing in one's own method, staying steadfast to the therapeutic project, and challenging the patient's problems according to an overarching strategy characterizes highly rated MBT. Paper III demonstrated satisfactory levels of initial working alliance among BPD patients in MBT irrespective of clinical outcomes and that a positive temporal development of alliance characterized treatments with good outcome. Focusing on tasks in therapy seems especially important among these patients. In terms of the overarching title, "Measuring MBT – a marriage of the common and specific psychotherapy factors", the thesis as a whole discusses how treatment processes in specialized treatment tailored for poorly functioning patients with BPD highlight how the CF and the specific factor approaches interact and suggests that "embedded alliance measures" (the alliance fostered by the specific technique) should be developed and implemented.

1.1 List of papers

Folmo, E. J., Karterud, S. W., Bremer, K., Walther, K. L., Kvarstein, E. H., & Pedersen, G. A. (2017). The design of the MBT-G adherence and quality scale. *Scandinavian Journal of Psychology*, *58*(4), 341–349. https://doi.org/10.1111/sjop.12375

Folmo, E. J., Karterud, S. W., Kongerslev, M. T., Kvarstein, E. H., & Stänicke, E. (2019). Battles of the comfort zone: modelling therapeutic strategy, alliance, and epistemic trust—a qualitative study of mentalization-based therapy for borderline personality disorder. *Journal of Contemporary Psychotherapy*, *49*(3), 141–151. https://doi.org/10.1007/s10879-018-09414-3

Folmo, E. J., Stänicke, E., Johansen, M. S., Pedersen, G., & Kvarstein, E. H. (2020). Development of therapeutic alliance in mentalization-based treatment—Goals, Bonds, and

Tasks in a specialized treatment for borderline personality disorder. *Psychotherapy Research,* 1–15. https://doi.org/10.1080/10503307.2020.1831097

1.2 Abbreviations

BPD: Borderline personality disorder

CF(s): Common factor(s)

DBT: Dialectical behavior therapy

EST: Empirically supported treatment

G-theory: Generalizability theory

GAF: Global Assessment of Functioning

IPA: Interpretative phenomenological analysis

MBT-G-AQS: The adherence and quality scale for mentalization-based group treatment

MBT-G: Mentalization-based group treatment

MBT-I-ACS: The adherence and competence scale for individual MBT

MBT-I: Individual mentalization-based treatment

PD(s): Personality disorder(s)

PDG: Psychodynamic group

RCT: Randomized controlled trial

RF: Reflective functioning

WAI-SR: The Working Alliance Inventory-Short Revised

1.3 Acknowledgements

On the first day of primary school, aged seven, I turned up with *Introduction to Psychology* by Harald Schjelderup under my arm. At the age of eight, I learned how to perform self-hypnosis. Although they did not encourage such maladaptive behavior, I must thank my mother and father for igniting such an interest in psychology and presenting me with the work of Sigmund Freud and Carl Gustav Jung, who were some of the first authors I discovered. In fact, my earliest memory is of my father reading to me when I was four years old, and my head sort of exploded as I came to understand that temperature was simply the movement of molecules. Now, at the age of 40, the time has apparently come for me to finish this Ph.D., which my friend Tor Espen Aspaas told me to pursue in 2002 (better late than never). At the finish line, the first of my next 20 "thanks" must go to Tor Espen for eternal kinship and for playing marvelous *Glass Bead Games* (Hermann Hesse) each week for many years—the rest is silence. Thanks to Andreas Lossius for being my friend ever since we met as medical

2 Introduction

2.1 Background

Clarkin and Levy (2004) describe the presence of personality disorders (PDs) in psychotherapy as follows:

> When an Axis II personality disorder is present, [therapists] should plan for disruptions in the treatment adherence and alliance. Many of the treatment manuals for symptom disorders such as anxiety and depression give insufficient information on approaches to patients with personality disorders who will present unique and difficult challenges in the treatment. (p. 202)

In the specialized and long-term treatments of borderline personality disorder (BPD), the development and maintenance of the therapeutic alliance constitutes a central issue of therapy and may constitute a central curing mechanism (Orlinsky et al., 2004; Spinhoven et al., 2007). However, it is well known that the therapeutic alliance can be extremely complicated in the treatment of BPD (Beck et al., 2015; Gabbard et al., 1988; Levy et al., 2010; Spinhoven et al., 2007; Yeomans et al., 2002; Yeomans et al., 1994; Young, 1999). According to Horvath (2001), "difficulties in developing an alliance with borderline and other personality-disorder clients had been identified in four studies" (p. 368), and there are indications that clients with poor alliance are more likely to drop out early in therapy (Horvath, 2001; Yeomans et al., 1994). Therefore, despite Falkenström et al. (2013a) reporting a six-fold stronger alliance effect with PD patients, "it is possible that the relatively weak overall relation between severity and strength of the alliance is, in part, due to the early attrition of more severe cases" (Horvath, 2001, p. 368). A systematic review of 41 studies of adults with BPD concluded that the overall completion rate was 71% for interventions of 12 months or longer and 75% for shorter interventions (Barnicot et al., 2011). However, these numbers may be lower due to a bias towards the publication of studies with higher completion rates (Barnicot et al., 2011). As the quality of alliance seems to depend on specific techniques in BPD treatment (Spinhoven et al., 2007), investigating the alliance in these treatments may bring more clarity regarding what works for whom (Roth & Fonagy, 2006) and the manner in which these treatments advance change. In structured, manualized BPD treatments, the quality of the alliance has been closely associated with therapeutic technique, and together the two factors have been found to facilitate positive outcomes (Folmo et al., 2019; Folmo et al., 2020; Kvarstein et al., 2020; Spinhoven et al., 2007). A clearer understanding of the alliance in evidence-based treatments for BPD may have considerable implications for clinical practice. The alliance has received

surprisingly little attention, despite relational pathology being at the core of BPD. This may be due to the current schism in psychotherapy research (Wampold & Imel, 2015) between those ascribing clinical change to a specific method and those who believe the common factors assumed by all traditions of psychotherapies to foster therapeutic change, are accountable for outcome. It is time to bridge this gap and find the common factors (CFs) between evidence-based treatments for BPD, as this may enhance the quality of treatment. Writing as a former BPD patient, Baltzersen (2021) states that the "study of personality disorders has come a long way, and this is characterized by the optimism prevalent within the community dedicated to its study" (p. 77). However, she concludes that "[d]elay in treatment and the gap between research and practice has severe consequences on both a personal and societal level. I challenge those studying PD to look to knowledge translation and implementation studies in developing strategies for change" (p. 79).

The main purpose of this dissertation is to investigate and measure mentalization-based treatment (MBT), one of the four major evidence-based treatments for BPD (Ellison, 2020). MBT is developed on the basis of traditional psychoanalysis and research on attachment and social cognition for treating BPD (Bateman & Fonagy, 2016; Karterud et al., 2020). It is a long-term, manualized, multicomponent, and psychodynamic treatment program (Bateman et al., 2012). In light of alliance research, it is of particular interest that a Norwegian study of MBT, which recruited patients from the same treatment cohort as the studies in the current thesis are based on, demonstrated low early drop-out rates (5%; Kvarstein et al., 2015). This thesis presents three articles but addresses four questions (the fourth being the overall topic of the theses). The three papers can be considered missing cornerstones in terms of defining, operationalizing, and identifying MBT (i.e., "measuring MBT"). The second purpose of the dissertation (particularly relevant for Paper II and Paper III) is to determine whether an investigation of MBT could simultaneously serve as an arena for bridging the division between those researchers attributing clinical change to specific techniques and those considering the CFs (between bona fide treatments) the curative aspect of psychotherapy (i.e., the major schism within psychotherapy; Wampold & Imel, 2015). The four investigations are interconnected and inform each other, but as this journey involves three different approaches to scientific knowledge, the introduction will outline the epistemological underpinnings essential for such an endeavor. Next, a description of the current state of psychotherapy research will serve as a prelude regarding how the current literature, interpreted through a potential marriage of the common and the specific factors, informs readers how best to

measure and understand MBT. Thus, the introduction zooms in from the abstract to the particular, meaning that the aims of the enunciated examinations emerge after an extensive exploration of psychotherapy research.

As the dissertation spans a vast realm of research often offering unclear, confusing, and conflicting findings, it may appear somewhat associative. Mathematics informs us that an associative property applies when the product is unaltered by the order of the factors. Such arithmetic seems iterated when entering the labyrinth of psychotherapy research. Consequently, the product of the dissertation may be an Ariadne's string; the thesis concludes that the term "embedded alliance", which was theoretically suggested by Hatcher (2010), empirically emerges like a promising path to pursue for future research, quality control, fidelity measures, and clinical supervision. Therefore, let me next summarize why I believe embedded alliance (the alliance fostered by the specific technique; Hatcher, 2010, p. 15) is crucial for measuring MBT and likely other psychotherapies. All concepts introduced in the paragraph below will be explained, elaborated, and empirically explored later.

2.2 Embedded alliance: A preluded conclusion

Contrasting interpretations of the meta-analytic results of psychotherapy led to a schism between those advocating the CFs, and those who believed that psychotherapy research ought to follow the example of medicine and pursue specific treatments that are efficacious in treating specific symptoms (Wampold & Imel, 2015). Only 17 of the 105 studies included in the meta-study by Flückiger et al. (2018) investigated CFs other than the working alliance, signaling its empirical dominance among the CFs of psychotherapy. However, there is no identifiable source or doctrine that owns the concept or can speak with authority concerning the alliance; it is a common factor because it exists by a consensus (Horvath, 2018). Consequently, the alliance is operationalized in research using an ever-increasing variety of assessment methods (Flückiger et al., 2018). The quality and nature of the alliance is shaped by the theoretical framework of the therapy (Bordin, 1979; Falkenström & Larsson, 2017). Consequently, it is "embedded" within the specific treatment method (or overarching strategy) applied (Hatcher, 2010). Mentalizing, the core ingredient in MBT, is the process by which we make sense of each other and ourselves, implicitly and explicitly, in terms of subjective states and mental processes (Fonagy et al., 2002). Before stating "that all mental health professionals will benefit from a thorough understanding of mentalizing" (Allen et al., 2008, p. 1), the conceivers of MBT argue that there are common factors, such as forming an

alliance, which is critical for successful psychotherapy. Hence, the architects of MBT may have suggested that mentalizing is "the most fundamental common factor among psychotherapeutic treatments" (Allen et al., 2008, p. 1), because "to be effective—in establishing a therapeutic alliance, for example—we clinicians must mentalize skillfully" (ibid.). However, placing mentalizing at core of psychotherapy, because it is part of creating an alliance, seems somewhat illogical—but that does not imply that mentalizing may not be a crucial common factor. In a study of the connections between common and specific therapeutic techniques by Tschacher et al. (2014), expert raters evaluated how strongly 22 CFs were implemented by the set of specific techniques in terms of the two dimensions "relevance" and "explanation". Raters indicated "relevance" and "explanation" for CFs such as "mitigation of social isolation", "provision of an explanatory scheme", "insight", "instillation of hope", "mindfulness", and "new narrative about self" (one may quite correctly suspect that most if not all the 22 CFs are highly correlated concepts). Patient engagement, affective experiencing, problem confrontation, and therapeutic alliance were judged most relevant for technique implementation. Mentalizing was rated high for "explanation" but low for "relevance", suggesting that mentalizing may not be the most relevant CF among psychotherapeutic treatments. As one would expect from the substantial allegiance effect reported within psychotherapy research, the ratings of relevance and explanatory power of the 22 different CFs, especially mentalizing, was associated with rater allegiance (Tschacher et al., 2014). Consequently, it is not surprising that mentalizing is championed as a crucial CF by the inventors of MBT. Mentalizing is a capacity intimately linked with normal (reflective) functioning, and an impairment in this ability is assumed to be associated with poorly functioning patients. This could partially explain why mentalizing was not rated the most relevant or explanatory technique by expert raters; mentalizing may not be the most relevant focus or CF in treatments for patients without personality pathology. In fact, a steadfast focus on mentalizing would perhaps leave such a client feeling they are being treated as a child, thus undermining trust. I will argue that the clear focus on mentalizing in MBT becomes a specific factor, despite mentalizing also being a CF. This may seem paradoxical, but the "common and specific factors address different aspects and levels of the psychotherapeutic process. It is therefore inadequate to contrast common factors with specific factors because these concepts reside at incommensurate logical levels" (Tschacher et al., 2014, p. 83).

Throughout this thesis, it should become abundantly clear that both CFs, such as therapeutic alliance, empathy, and goal consensus, and specific techniques, such as exposure, empty-chair

technique, paradoxical intention, and particular forms of interpretations, have been "found to be consistently and strongly associated with positive therapeutic outcome" (Orlinsky et al., 2004, p. 341). Consequently, the question becomes how the common and specific factors can be successfully tailored to a specific patient (Norcross & Wampold, 2011). In terms of MBT, the term "embedded alliance" may help us spell out how mentalizing is intertwined with alliance, which is after all considered the quintessential CF in psychotherapy research. I believe that embedded alliance should address working alliance (e.g., agreement on tasks and goals) in line with the therapeutic model (theory of pathology and change). A deep understanding of the etiology of BPD seems necessary to define embedded alliance in evidence-based treatments for BPD, such as MBT. According to MBT theory, BPD patients are handicapped in terms of mentalizing capacity due to factors such as inadequate mirroring and a lack of trust in their own inner reality or mentalizing abilities. Therefore, an embedded alliance in MBT would, for instance, target the unalignment between thoughts and feelings typical for BPD patients. Mentalizing would be the core process necessary to bridge this gap because the normative maps of social reality are not typically expressed verbally (Bateman & Fonagy, 2016; Karterud et al., 2020). Mentalizing can partially be described as verbalizing normative social rules and learning the process of applying these rules, which are typically expressed non-verbally. While verbalizing implicit social rules to well-functioning clients would be belittling them, it may prove essential to correct the entrenched misalignment with such social norms typically caused by childhood trauma or other factors associated with the development of BPD pathology (Gullestad & Wilberg, 2011). Therefore, in MBT, the embedded alliance would be exemplified by the therapist making one's own inner reality transparent in parallel to making the patient aware of their inner reality (e.g., denoting feelings, validating, challenging beliefs, and acknowledging mentalizing). This mentalizing process would possibly make the implicit ingredients—the necessary mirroring of the social components the BPD patients' environment typically failed to foster (e.g., marked mirroring—explicit. We can see a somewhat similar process (unveiling implicit social norms) in the learning of skills in dialectical behavior therapy (DBT; Linehan, 2014), and the role of the MBT therapist as a teacher of psychological content—through a mentalizing discourse—seems of paramount importance. Consequently, the thesis will also debate the art and aspect of pedagogical stance, recently proposed as an MBT intervention (Karterud et al., 2020). Let us unfold this preluded conclusion. I will start by providing some necessary color to our overall map of psychotherapy research.

2.3 The great puzzle of psychotherapy research

It has been demonstrated that the average effect size of psychotherapy is about 0.8 (Wampold & Imel, 2015; Wampold et al., 2007b), which means that about three-quarters of patients who receive psychotherapy are better off than those left to recover by themselves (Fonagy, 2010; Roth & Fonagy, 2006). However, there is a great puzzle in psychotherapy research concerning the effective components of psychotherapy. Although it is well established that individual psychotherapy is effective, dozens of meta-analyses of psychotherapy with widely varying hypotheses about change have not shown one type of therapy to be more effective than another. Furthermore, with the exception of the alliance, no robust mechanisms of change have been identified (Wampold & Imel, 2015). Interestingly, in the research on BPD there is solid evidence that various specific therapies are superior to treatment as usual (Ellison, 2020). These evidence-based treatments have many elements in common (e.g., highly structured programs) that may be partly responsible for their effects, but there is no reliable evidence that any of these specific treatments is more effective than any other (ibid.). Thus, before embarking on the task of investigating (measuring) in depth any specific type of psychotherapy for BPD, which is the aim of the current dissertation, it is imperative to take one step back and explore if and why this would be a good idea in light of both current empirical data and the main competing theories trying to account for these findings. As this is also one of the most central debates in current psychotherapy research, such an inquiry should also be of interest to a broader audience.

Jules Henri Poincaré wrote that "Science is built up with facts, as a house is with stone. But a collection of facts is no more a science than a heap of stones is a house" (Aitchison, 2000, p. 12). Therefore, the collection of facts must be digested in a meaningful way—the "lion is made of assimilated sheep" (Valéry, 2015, p. 10)—and put into the context of an overarching theory/model. Apparently, there are many ways to digest the same thing; for example, "when a venomous snake drinks water, it becomes poison; when a cow drinks water, it becomes milk" (Zen proverb). Observing the same facts, psychotherapy researchers are divided in their interpretations (Wampold & Imel, 2015). Some believe that focusing therapy and therapy training on the so-called CFs is most effective, while others champion specific factors, that is, the empirically supported treatment (EST) approach. Before we turn our attention to how we think these interpretations differ and explore whether these conflicting views can be illuminated and/or differentiated by, for instance, applying Lakatos' (1970) Methodology of

Scientific Research Programmes to psychotherapy research, I will first provide the epistemological framework for this dissertation.

2.4 Wilber's four quadrants as an epistemological backbone

Habermas' (1986) distinction between general theories and general interpretations and the subdivision of validity claims into objective truth (It), subjective sincerity (I), and intersubjective justness (We) is the epistemological foundation for this thesis. It therefore uses three different strategies (validity claims) to search for knowledge. Wilber (2001) states that "Habermas's three validity claims, for truth (objects), truthfulness or sincerity (subjects), and rightness or justice (intersubjectivity), refer respectively to the Right half, the Upper Left, and the Lower Left" (p. 149). Wilber (2001b) evolved this theory by adding a fourth validity claim denoted "Its/They", resulting in the following four quadrants (see Figure 1):

1) Interior individual perspective (upper-left quadrant) would be the science of literature but also Freudian psychoanalysis, which interprets people's interior experiences and focuses on "I".

2) Interior plural perspective (lower-left quadrant) would, for instance, include Gadamer's philosophical hermeneutics, which interpret the collective consciousness of a society or plurality of people and focus on "We".

3) Exterior individual perspective (upper-right quadrant), which can be exemplified by B.F. Skinner's behaviorism that treats the internal experience of the subject as a black box and considers the subject a specimen to examine, an "It".

4) Exterior plural perspective (lower-right quadrant), which includes Marxist economic theory that focuses on the behaviors of a society (i.e., a plurality of people) as functional entities seen from outside, or "They".

The theory of the four quadrants implies that one needs different methods within different knowledge domains (validity claims) and informs us that the common denominator between different scientific approaches is not that it searches for or finds objective truth but that it is transparent in its method to build theories from systematic observations (Wilber, 2000b). The scientific study of literature and art relies on the researcher assimilating the relevant culture(s) and being sincere in their interpretation of the work. It seems that one of the major reasons for the current state of psychotherapy research is that the scientific study of psychology has tried to approach an "ideal objective science" (upper-right quadrant), which has favored the medical model, largely supported by randomized controlled trials (RCTs) (Wampold & Imel,

2015). One of the most successful theories categorized within this validity criterion (It) is quantum theory (i.e., quantum electrodynamics), which describes the "real world" with a precision "equivalent to measuring the distance from Los Angeles to New York to the thickness of a human hair" (Feynman, 2006, p. 118). "Medicine is of all the Arts the most noble" according to *The Hippocratic Corpus*, and approximately 2300 years later the field of medicine still accounts for a substantial number of impressive findings within the upper-right quadrant. For example, one recent nested case-control study of 58,769 patients with a diagnosis of dementia and 225,574 matched controls found that anticholinergic drugs—which work by regulating muscle contraction and relaxation—may increase a person's risk of developing dementia (Coupland et al., 2019). It would be great if psychotherapy could compete with such results within this quadrant (It; validity criterion), but seeing that Luborsky et al. (1999) reported that researcher allegiance in psychotherapy research has been found to account for 70% of the variance in the effect sizes of treatment comparisons, it seems reasonable to agree with Messer and Wampold (2002), who state "How odd it is, then, that we continue to examine the effect of different treatments (accounting for less than 1% of the variance) when a factor such as the allegiance of the researcher accounts for nearly 70% of the variance!" (p. 23). Therefore, one may wonder whether progress within psychotherapy research may in fact necessitate a return to case studies or other scientific approaches more linked to subjective sincerity (I/upper-left quadrant), that is, qualitative research. The current thesis tries to incorporate three of the quadrants ("I", "We", and "It"); for example, Paper II applied philosophical hermeneutics to model concepts such as alliance and strategy in four case studies. Before investigating this further (this topic runs as a thread through the entire dissertation), let us first consider the current standing of psychotherapy research, which will inform us what would be most meaningful to measure in MBT.

2.5 The great psychotherapy debate

Perhaps the most striking schism in psychotherapy research is the opposition/dichotomy between focusing on so-called (CF; Frank & Frank, 1993; Wampold & Imel, 2015) and specific factors, that is, the EST approach. Non-specific elements in psychotherapy refer to elements of therapy that are shared across virtually all *bona fide* treatments. They have been said to include a healing setting, education, a treatment rationale, expectations of improvement, a treatment ritual, and the therapeutic relationship (Wampold & Imel, 2015). Frank and Frank (1993) suggest that all treatment methods share several non-specific elements, while each may have its own specific ingredients. Specific factors refer to the

technical interventions therapists engage in based on their theoretical orientation (Butler & Strupp, 1986). CFs and alliance predict more variance in outcome than specific factors (Wampold & Imel, 2015). However, the examination of therapeutic factors other than treatment methods has been considered "unscientific":

> Of all the aspects of psychotherapy that influence outcome, the treatment method is the only aspect in which psychotherapists can be trained, it is the only aspect that can be manipulated in a clinical experiment to test its worth, and, if proven valuable, it is the only aspect that can be disseminated to other psychotherapists. (Chambless & Crits-Christoph, 2006, p. 199)

2.6 The common factor approach

The CF approach (Frank & Frank, 1993; Wampold & Imel, 2015) conceptualizes psychotherapy as a socially constructed and facilitated healing tradition. According to Wampold and Imel (2015), the CF model focuses on five elements that are necessary and sufficient for change: (1) an entrusting curative setting in which treatment takes place, (2) the therapist provides a rationale for the therapy, which is accepted by the patient, (3) a culturally embedded explanation for the psychological disorder or psychic distress, (4) an emotional bond between the therapist and patient, and (5) a therapeutic ritual/procedure that promotes positive and progressive behavior. Laska et al. (2014) state that the EST approach prioritizes different theories of change, while the "CF approach states that the adoption of a credible theory is only one aspect of many necessary common factors that contribute to behaviour change" (p. 469). EST follows a few core assumptions from the CF approach. The first assumption is that any therapy that contains *all* five elements of the CF approach will be efficacious for the presenting problem being treated (ibid.). A second assumption is that factors of the therapeutic relationship (e.g., the working alliance) should foretell the outcome of psychotherapy. A consequence of this is that there will be differences among therapists; that is, more effective or "talented" therapists will more skillfully provide such a treatment or "magic potion" (Fonagy, 2010). A third assumption is that treatments intended to be therapeutic will be superior to "supportive control" or psychological placebo conditions. Notably, several researchers have challenged the CF perspective: "Our view is that the "CF perspective" should be subject to the same sorts of empirical investigations as any other "perspective" on behaviour change" (Crits-Christoph et al., 2014, p. 491). This approach has also been criticized for attempting reverse engineering:

We are concerned that the CF approach will not make rapid progress because it appears to rely on reverse engineering . . . it attempts to extract core therapeutic strategies by inferring and inducting them from a heterogeneous set of outcomes gathered across innumerable studies, patient groups, intervention intensities and durations, and so on. (Baker & McFall, 2014, p. 484)

2.7 Empirically supported treatment

The goal in the EST approach is to establish a causal relationship between treatment and outcome (Fonagy, 2010):

Templar Knights of psychotherapy, who were both fearsome warriors and devout monks, could righteously believe that the people they cured recovered because of their magic spells and carefully measured potions, which often took decades of apprenticeship to learn to brew with confidence. (Fonagy, 2010, p. 22)

To qualify as an EST, a treatment should typically adhere to the following three standards (Benjamin & Critchfield, 2010; Kazdin, 2011): (1) the treatment can be defined and taught by a manual, (2) a superior effect for a specific diagnosis in an RCT has been shown (with an adequate control group), and (3) such results should be replicated at more than one site (Chambless & Ollendick, 2001). Within the EST approach, the therapeutic relationship (e.g., the working alliance) is not seen as a curative ingredient in itself but as an implicitly supporting factor. However, there have been attempts to manualize the therapy relationship itself (e.g., Crits-Christoph et al., 2006a), but no manualized set of relational interventions has passed the p < .05 test in RCTs, and "effective relating" is not recognized as a change agent or an EST (Benjamin & Critchfield, 2010). There are two core EST assumptions: (1) treatment specificity and (2) disorder specificity. Therefore, each EST advocates a specific mechanism of change based on a given scientific theory; Barlow (2004) writes that ESTs contain "specific psychological procedures targeted at the psychopathology at hand" (p. 873). The evidence-based movement emphasizes the empirical demonstration (RCT designs) of specific therapies' effectiveness in the treatment of particular disorders (Budd & Hughes, 2009; Messer & Wampold, 2002; Norcross & Goldfried, 2005; Stiles et al., 1986). However, even when a therapy has been shown to be responsible for change in general (RCT), other factors may cause apparent reported changes. RCTs have been criticized for often having poor statistical power and poor generalizability (e.g., as a result of restricted sample sizes) and for being "causally empty" (Cook et al., 1979; Elliott, 2002; Haaga & Stiles, 2000; Kazdin, 1998). Two ESTs could easily have the same fundamental mechanism of action in reality,

despite contradictory theory (i.e., the theory is imprecise). Regarding RCTs, one should also be cautious when "concluding that patients with a certain diagnosis should always be offered a certain specific treatment" (Jørgensen, 2019, p. 54). For such reasons, it has been argued that inferring causality requires another condition. The provision of a logical mechanism or possible causal relationship (e.g., Haynes & O'Brien, 2003)—that is, science should be driven by theory (Elliott, 2010)—therefore advances change process research (CPR) as a critical complement to RCT and argues that single-case student research projects are a feasible alternative to qualitative interview research in therapeutic training programs.

> By staying close to clinical practice, significant events studies can appeal to practice-oriented students in many of the same ways that qualitative interview studies do, while actually being more grounded in practice by virtue of exposing students to actual therapeutic practice as opposed to talk about practice. (Haynes & O'Brien, 2003, p. 131)

2.8 Common factors versus specific ingredients

Two major areas of research have been used as evidence that non-specific factors are the principal mechanisms of change in psychotherapy: 1) the evidence of outcome equivalence between two or more psychotherapies and 2) the finding that alliance is the most and only robust mechanism of change (Wampold & Imel, 2015). Nevertheless, we do not know whether the alliance is mostly due to the therapist, the patient, their interaction (in the statistical sense), or symptom improvement. It may just be that good clients form good alliances and are destined to get better (Barber et al., 2000; DeRubeis & Feeley, 1990; Garfield, 1978; Gibbons et al., 2003; Martin et al., 2000). Further, DeRubeis et al. (2005) argue that outcome equivalence between different treatments in RCTs does not imply that the same mechanisms produce the outcomes. Laska and Wampold (2014b) state that

> We need to reiterate that there is no such thing as a "common factor" treatment. One of the aspects of all treatments is that the patients are provided an explanation for their disorder and that there are treatment actions consistent with that explanation. That is, the psychotherapy offered to the patient must contain a cogent explanation for the patient's distress and a plan for overcoming his or her problems. (p. 520)

2.9 The working alliance

As psychotherapy research found no significant differences in terms of efficacy (between bona fide psychotherapies), the search for the effective "ingredients" common to different

kinds of treatments was initiated: The "common factors." (Horvath, 2018, p. 501). As the importance of the relationship between the patient and the healer has been recognized since ancient times, "the helper–client relationship was identified as an obvious candidate" (Horvath, 2018, pp. 501–502). The most researched aspect of the helper–client relationship is the alliance. However, as the concept of the alliance exists as a consensus, it is not constrained within a theoretical framework (Horvath, 2018, p. 500), and an interesting question becomes which relational aspects the alliance includes—for instance whether the real relationship (RR) is part of the alliance or not. Bordin suggested that the RR was captured by the bond aspect of the alliance (Bordin, 1994), while (Gelso & Carter, 1985) theorized that the therapy relationship has three components: Transference, alliance, and the RR. Relational elements such as alliance, empathy, warmth, trust, and genuineness represent different levels of abstractions, and "there is a practical vacuum in the literature addressing questions about the relations among these elements, both from the conceptual and from the empirical perspective: How much do these elements overlap or contribute to each other?" (Horvath, 2018, p. 511).

The alliance is considered one of the five elements necessary for therapeutic change (Wampold & Imel, 2015) in the CF approach. However, the proponents of the CFs have no copyright on the importance of the therapeutic relationship. According to Rogers (1951), the three effective components of psychotherapy are empathy, unconditional positive regard, and congruence. The psychoanalyst Greenson (1965) claims that a working alliance is as important as analyzing the patients (e.g., the transference neurosis). Luborsky (1976) applies a counting signs method of assessing alliance and describes two types of alliance, one "based on the patient's experiencing the therapist as supportive and helpful" and one "based on a sense of working together in a joint struggle" (p. 94). Bordin (1979) redefines the working alliance in terms of a collaboration between therapist and patient while engaging in a series of tasks tailored to lead toward agreed-upon goals. Parallel to that process, a bond develops that supports the patient's capacity for positive and trustful states. Bordin (1979) claims that across therapies "the effectiveness [is] in part, if not entirely, a function of the strength of the alliance" (p. 253). A variety of definitions of alliance and relationship have shown robust associations with treatment outcome (Norcross et al., 2006). However, Horvath (2006) calls for "a clearer definition of the alliance", a "consensus about the alliance's relation to other elements in the therapeutic relationship", and clarification of "the role and function of the

alliance in different phases of treatment" (p. 258). Throughout this thesis, we will elaborate the view that "embedded alliance" is a comforting term when battling with such questions.

The definition of the therapeutic alliance proposed by Bordin (1979) was redefined by Horvath and Luborsky (1993) as a "pan-theoretical concept". Bordin's formulation highlights the collaborative relationship between patient and therapist in the common quest to overcome the patient's suffering and (self-) destructive behavior and consists of three essential elements—(1) agreement on the goals of the treatment, (2) agreement on the tasks, and (3) the development of a personal bond made up of reciprocal positive feelings. The variable influence of the alliance in different therapies has led some to propose that this relationship may play differing roles across treatment modalities (Gaston et al., 1998; Safran & Wallner, 1991). In terms of the therapeutic relationship (e.g., the working alliance), correlation studies show that alliance (Bordin, 1979, 1983, 1994; Gaston, 1990; Luborsky, 1976) at the onset of treatment predicts improvement in symptoms at the termination of treatment (Barber et al., 2000; Cloitre et al., 2004; Horvath & Bedi, 2002; Horvath & Symonds, 1991; Klein et al., 2003; Martin et al., 2000). But as we all know, establishing correlation is but an illusion of explanation (Barber, 2009), and even though meta-analyses from 2011 and 2018 (Flückiger et al., 2018; Horvath et al., 2011) report an aggregate association between alliance and outcome of .275–.78 (typically in the range of .20–.30), there remains an animated argument about the curative potency of the alliance, predominantly in therapies examined using RCT designs (Horvath et al., 2011; Ulvenes et al., 2012b). Notably, "therapists' individual differences have been found to predict alliance quality and treatment success" (Muran et al., 2010, p. 321). Therefore, Fonagy (2010) and Lemma et al. (2011) conclude "So the ability to form an alliance does mark out our more talented therapists, but what it is that they do more or less of that makes them more or less effective still remains a mystery" (p. 37; p. 17).

Despite the assumption that it takes the devoted apprentice decades to learn how to brew these carefully measured psychotherapeutic potions (Fonagy, 2010), we know little about how "talented therapists" obtain their recipes and administering abilities, as the current literature seems to exclude major effects of therapist training (Beutler, 2004; Miller & Binder, 2002; Ogles et al., 1999; Rønnestad & Ladany, 2006): "Overall, these findings tend to cast doubt on the validity of the suggestions that specific training in psychotherapy, even when unconfounded with general experience, may be related to therapeutic success or skill" (Beutler, 2004, p. 239). This is regarded as true for training in manualized short-term

psychodynamic treatment (Bein et al., 2000). However, there are some indications for further research; for example, Crits-Christoph et al. (2006a) have reported non-significant but promising results from training five therapists (45 patients) in alliance-fostering techniques.

Applying a somewhat seemingly circular logic, Wampold and Imel (2015) simply state that the most important common curative factor is treatment itself, while others call for studies on how psychotherapy actually leads to change (e.g., Elliott, 2011; Greenberg, 2007; Kazdin, 2009). Elliott (2011) emphasizes that even though there are many theories about therapeutic change, we know little of how change actually occurs. Kazdin (2009) concludes that "[a]fter decades of psychotherapy research and thousands of studies, there is no evidence-based explanation of how or why even the most well-studied interventions produce change, that is, the mechanisms through which treatments operate" (p. 426). The foremost enigma today is the last of the four questions Klaus Grawe articulated in 1997: How does psychotherapy work?

2.10   Methodological differences between the two approaches (EST and CF)

Although RCT designs have demonstrated that psychotherapy has an effect compared to no treatments, these studies account for surprisingly little comparative outcome variance (Rønnestad et al., 2006; Wampold & Brown, 2005; Wampold & Imel, 2015). An RCT design typically accounts for about 1% (Cohen's d of 0.20) of the outcome variance for differences between treatments (Messer & Wampold, 2002) and predicts about 5%–7.5% of the variance in outcomes attributed to the alliance (Baldwin et al., 2007; Flückiger et al., 2018; Horvath & Bedi, 2002; Horvath et al., 2011; Martin et al., 2000; Wampold & Imel, 2015). Variability in outcome explained by specific ingredients accounts for roughly 0%: "That is, the preponderance of effects were near zero and the frequency of larger effects was consistent with what would be produced by chance, given the sampling distribution of effect sizes" (Wampold & Imel, 2015, p. 133). In comparing active treatments (in contrast to the more usual comparisons of active treatments with controls), Luborsky et al. (2002) found a non-significant effect size of 0.20 based on 17 meta-analyses, shrinking down to 0.12 when corrected for researcher allegiance to different treatments. In line with such findings, Grissom (1996) meta-analyzed 32 meta-analyses of comparative treatments and reported an effect size of .23, while Wampold et al. (1997) "found an effect size identical to that of Luborsky et al., namely, .20" (Messer & Wampold, 2002, p. 21); "an equivalent Pearson's r would be .10" (Luborsky et al., 2002, p. 2). Such meta-analyses demonstrating equivalence between

different treatment methods fail to establish causality between treatment and outcome (Fonagy, 2010) and cast doubt on the power of the medical model of psychotherapy.

Discrepancies in psychotherapy research stem from methodological challenges, multiple variables to measure outcome, inappropriate control conditions in RCTs, statistical controversies, and biased reporting of results (Elliott, 2002; Fonagy, 2010; Luborsky et al., 1999; Wampold et al., 2017). For example, as Wampold and Bolt (2006) state, "We demonstrate that Elkin et al. [2006] chose a model and performed various operations that increased the likelihood that therapist effects will be absent" (p. 184). Messer and Wampold (2002) therefore emphasize the CFs and therapist effects and advocate the following for psychotherapy research and practice: 1) "Limit clinical trials comparing bona fide therapies because such trials have largely run their course. We know what the outcomes will be." 2) "Focus on aspects of treatment that can explain the general effects or the unexplained variance in outcomes." 3) "Cease the unwarranted emphasis on ESTs. They are based on the medical model, which has been found wanting and wrongly leads to the discrediting of experiential, dynamic, family, and other such treatments" (p. 23–24).

As prophesized by the CF approach, in studies examining the effectiveness of different therapy methods, one finds more variation within each method than between them, presumably due to the differences among therapists (Wampold & Imel, 2015, p. 259). Research suggests that some therapists are consistently more helpful than others and that these same therapists are better able to facilitate the development of the therapeutic alliance (e.g., Baldwin et al., 2007; Luborsky et al., 1986; Wampold & Imel, 2015). Nissen-Lie et al. (2010) state that "[a]ccordingly, a more justified, alternative conception to 'the dodo bird verdict' […] could be that some therapists win and some do not, independent of the therapeutic method they use" (p. 627). This can be considered an argument for the appropriateness of studying the technique of expert therapists (Lambert & Ogles, 2004). However, some studies indicate no therapist effects in RCT trials (Elkin et al., 2006), while others report rather substantial therapist effects, explaining a range from 0.2% (Wampold & Brown, 2005) to 29% of the variance in outcomes (Baldwin & Imel, 2013; Baldwin et al., 2007; Blatt et al., 1996; Dinger et al., 2008; Huppert et al., 2001; Johns et al., 2019; Lambert, 1989; Lutz et al., 2007; Orlinsky & Howard, 1980; Wampold, 2010; Wampold & Bolt, 2006; Wampold & Brown, 2005; Zuroff et al., 2010). In a systematic review of therapist effects (Johns et al., 2019), the weighted average was 5% (8.2% in RCTs) with a range from 0.2%–

29%. Notwithstanding the impressive inconsistency in this reported range, knowing that the therapist explains a considerable portion of the variance in outcome (Wampold, 2010) is not an explanation of the factors separating helpful from less helpful therapists (Beutler, 2004). However, in an investigation of therapist effects on outcome and alliance in inpatient psychotherapy, the therapists accounted for a much greater variability in alliance (33%) than in outcome (3%) (Dinger et al., 2008).

For the above outlined reasons, both passionate potion brewers and psychotherapy researchers have requested methods that can shed light on what actually accounts for therapeutic change (Elliott, 2002, 2010; Kazdin, 2005, 2009; Lemma et al., 2011). This could be seen as an argument for the aptness of qualitative methods (upper-left quadrant). Single-case experimental designs have been suggested as one way to introduce systematic assessment and evaluation in clinical practice (e.g., Barlow et al., 2007; Kazdin & Tuma, 1982; McLeod, 2011), and there has indeed been increased interest in the study of individual instances (e.g., Busse et al., 1995; Elliott, 2002, 2010; Hilliard, 1993; Iwakabe & Gazzola, 2009). Stiles (2005, 2007, 2009) argues that case studies are an important supplement to group-level statistical hypothesis testing, where unique features most often are considered as errors (Rosenwald, 1988), and can point out where theories need to evolve: "If you restrict yourself to the themes that are common across cases, you will overlook the most interesting parts. Each case tells us something new, and new observations are always valuable, whether they confirm previous theory or add something unexpected" (Stiles, 2007, p. 123). Consequently, an adequate theory has to encompass the unique qualities of each case, as well as the common features (Elliott, 2002; Habermas, 1986).

2.11 Proliferation by rivalry of theories

To make sense of the two main competing theories/approaches/research programs (the CF approach and the EST approach) in psychotherapy, we need to focus on what exactly is being delivered in these treatments, what are in fact the ingredients in these "magic potions" (Fonagy, 2010). We therefore use adherence and competence in the present thesis to disclose similarities and differences between these two strands of psychotherapy/process research. Applying Lakatos' theory (Methodology of Scientific Research Programmes), I argue that the apparent dichotomy/conflict between these major programs is in fact not clear at all but that each approach actually may be dependent on the other as an auxiliary hypothesis. That is, CFs do not exist without a specific treatment method; "there is no such thing as a common factor

treatment" (Laska & Wampold, 2014a, p. 520). This is in line with the scientific methodology of Lakatos, who states that "The problem fever of science is raised by the proliferation of rival theories rather than by counterexamples or anomalies. This shows that the slogan of proliferation of theories is much more important for sophisticated than for naïve falsificationism" (Lakatos, 1970, p. 121). It should be noted that in their article titled "Common factors as a therapeutic approach: What is required?", Bailey and Ogles (2019) propose a theory of change, including a theory of pathology, rationale for addressing pathology using CF that can be identified with a specified rationale, and therapeutic practices to evoke the identified CF. We find a therapeutic approach based on CF conceptually troublesome to categorize; as Wittgenstein (1975) states in the opening of his first lecture on the foundations of mathematics at Cambridge in 1939, "[W]hen we learn spelling, we learn the spelling of the word 'spelling' but we do not call that 'spelling of the second order'" (p. 14). Therefore, my view is that specifying CFs as a therapeutic approach would be a specific technique (EST), which like other specific psychotherapies cannot exist without the CF.

2.12 Lakatos' Methodology of Scientific Research Programmes

In defense of the rationality of science (Benton & Craib, 2010, p. 62), Lakatos states that

> A 'model' is a set of initial conditions (possibly together with some of the observational theories) which one knows is bound to be replaced during the further development of the programme, and one even knows, more or less, how. This shows once more how irrelevant 'refutations' of any specific variant are in a research programme: their existence is fully expected, the positive heuristic is there as the strategy both for predicting (producing) and digesting them. (Harding, 1975, p. 244)

Further, what Lakatos denotes "naïve falsificationism" holds that a bundle of hypotheses (i.e., a hypothesis/core theory and its background assumptions/auxiliary hypotheses) as a whole can be tested against the empirical world and be falsified if it fails the test. The Duhem–Quine thesis (after Pierre Duhem and Willard Van Orman Quine) says it is impossible to isolate a single hypothesis in the bundle. This philosophical approach (Lakatos) allows for a more pragmatically tailored backbone for psychotherapy research:

> Practice has to tell researchers where knowledge is most needed and to ensure that science is firmly grounded in everyday clinical care. Best evidence is only meaningful if used in proper argumentation. Argumentation is only meaningful if based on the best evidence in its building blocks. (Fonagy, 2010, p. 38)

2.13 Popper, Kuhn, and Lakatos: Making sense of psychotherapy evidence

The principal criticism of the CF approach is that it is tautological, unfalsifiable, and therefore not subject to the same scientific rules as (ESTs; Laska & Wampold, 2014b). However, empirical observations are arguably inherently theory-loaded. "Scientists see new and different things when looking with familiar instruments in places they have looked before" (Kuhn, 1996, p. 111). This is in opposition to Popper's (1980) idea that a good scientific theory is "guesses [...] guided by the unscientific" (Popper, 2005, p. 278). He states that empirical data are "irrelevant to the logical analysis of scientific knowledge" (Popper, 2005, p. 7).

> Thus Popper tends to describe theories as 'guesses about the structure of the world' (C.R., p. q5), and answers the question 'How do we jump from an observation statement to a good theory?'–'. . . by jumping first to any theory and testing it, to find whether it is good or not' (C.R., p. 55). (Cosin et al., 1971, p. 125)

In line with Kuhn's view, Laska et al. (2014) state

> that restricting the lens through which a phenomenon is examined (in this case, psychotherapy) restricts what can be observed and the manner in which 'evidence' is interpreted. In other words, a restricted scientific aperture means less of the evidentiary picture is in focus. (p. 468)

Lakatos bridges the ideas of Kuhn (scientific revolutions) and Popper (critical rationalism) and is therefore well suited for psychotherapy research. Wampold and Imel (2015) state that "Simply put, Lakatos presents a reconstruction, which is eminently useful in making sense of psychotherapy evidence" (p. 63).

2.14 Two rival research programs: CF and EST

Research projects are guided by presumptions about the phenomena that inform the eventual empirical findings. According to Kuhn (1970), when the scientist is occupied with a research problem, they must premise current theory as "the rules of his game." Lakatos (1970) states that "the choice of a theory is equally the choice of a research programme" (p. 262). Lakatos' term "research program" closely resembles Kuhn's notion of "paradigm" in that it defines the activities of the researcher in the field concerned (Benton & Craib, 2010, p. 61). However, Lakatos' research program consists of a hard core that contains the basic assumptions and main tenets of the theory and a "protective belt", a surrounding defensive set of *ad hoc* hypotheses. According to Lakatos, the auxiliary hypotheses of the "protective belt" can be amended to fit observations as long as the "hard core" is untouched. If these changes in the

auxiliary hypotheses foster improved predictions or explanatory power, the program is *progressive*; if the opposite is true, it is termed *degenerative*. According to Lakatos, a theory ("hard core") should not be abandoned unless there is another progressive alternative. Wampold and Imel (2015, p. 67) argue that there are two main rival research programs in the field—the Contextual Model (CF), which they reason is progressive (Lakatos, 1970), and the Medical Model (EST), which they claim is degenerative.

2.15   Potential marriage of EST and CF

Falsification cannot "compel the theorist to search for a better theory" simply because falsification cannot precede the better theory (Lakatos et al., 1980, p. 37). Consequently, ad hoc adjustments can produce a richer theory in the sense that the revised theory is better able to explain how the phenomenon works (in Lakatos' terms, the research program is progressive). However, the ad hoc adjustments can become arduous, as the theory becomes burdened with amendments and is found wanting, particularly if an alternative theory exists that can explain observations more parsimoniously and can anticipate observations under various conditions. As noted above, Wampold and Imel (2015) argue that the EST should be considered a degenerative research program and therefore be abandoned. To prelude our discussion, in line with Folmo and Langjord (2014), this conclusion seems premature and futile for future research. Even though there is converging evidence for a statistically non-significant difference between bona fide psychotherapies, this does not mean that there are no clearly identifiable differences between these treatments and that their "magic potions" (Fonagy, 2010) contain specific ingredients that are, for instance, moderated by CFs. For example, Ulvenes et al. (2012b) analyzed 46 psychotherapy sessions from a previous RCT (Svartberg et al., 2005) and argue in line with other recent studies (Hoffart et al., 2012; Wampold & Budge, 2012) that alliance has an indirect effect, such that it is necessary for other factors, such as specific ingredients, to work. Lakatos stresses that "[f]or the sophisticated falsificationist proliferation of theories cannot wait until the accepted theories are "refuted" (or until their proponents get into a Kuhnian crisis of confidence)" (Lakatos et al., 1980, p. 37). However, as will be returned to in the discussion, this could signal that the rivalry or perhaps merger of these two research programs might be more in line with this philosophical underpinning.

2.16   Clinical significance and treatment efficacy

As this dissertation examines an evidence-based treatment for a specific disorder, it becomes of principal importance to investigate what effectiveness actually means. Is it a statistical term based on, for instance, aggregated symptom reduction, and/or is it improvement that appears meaningful for the individual and the social environment? Jacobson and Truax (1991) criticize Smith et al. (1980) for confusing statistical effect or effect size with efficacy and introduced the term clinical significance. The clinical significance of a treatment refers to its ability to meet standards of efficacy set by the government, patients, therapists, and researchers (Jacobson & Truax, 1991; Lambert & Ogles, 2004). In contrast to criteria based on statistical significance, judgments concerning clinical significance are based on external standards specified by interested parties in society. Numerous suggestions of criteria for operationalizing clinical significance have been put forth. Kazdin (1978) proposes that behavior changes should be viewed as clinically important if the intervention brought the client's performance within the range of socially acceptable levels, as evidenced by the client's peer group, or if the client's behavior is judged by others as reflecting a qualitative improvement on global ratings. Lambert and Ogles (2009) recommend that an estimate of clinical significance based on Jacobson and Truax (1991) is included in all psychotherapy outcome studies. Today, within the clinical population, a return to normal functioning (40–60% of patients typically return to normal functioning; Lambert & Ogles, 2004) is a consensus standard for treatment efficacy, but as mentioned above there are considerable methodological problems concerning outcome measures (Fonagy, 2010). According to Kazdin (2006), measures for the most part are arbitrary, measuring subtle psychological processes on arbitrary scales. Yet, we reify them; we treat and think of them as if they resemble something self-evident in the outside world. Lemma et al. (2011) ask "[w]hat is the real life significance of change of 0.5 on the GSI score of the SCL-90?" (p. 15). This points to the importance of investigating the reliability and validity of such scales and measures.

Observing that around half of patients recover during psychotherapy implies that a relatively large proportion of patients do not benefit from the treatments they receive (often denoted non-responders). Linden and Schermuly-Haupt (2014) found an emerging consensus that unwanted events should be expected in 5%–20% of psychotherapy treatments. Lambert and Ogles (2004) report that 15%–20% of patients show no significant change, while 5%–10% consistently deteriorate during the course of treatment (e.g., Crawford et al., 2016; Jarrett, 2008; Lambert, 2013; Mohr, 1995). Unfortunately, there has been limited research on patients who do not respond to psychotherapy. However, there are indications that the likelihood of

non-response or negative response increases "with more severe symptoms, with more profound functional impairment, with more problems in interpersonal relatedness, and with the presence of personality disorders" (Solbakken & Abbass, 2014, p. 2). Patients with comorbid personality disorders (PDs) have been found especially challenging and resistant to treatment, resulting in a more negative process, higher attrition rates, and increased treatment length (Benjamin & Karpiak, 2001; Clarkin & Levy, 2004; Westen & Morrison, 2001). The latter may be understood as an argument that long-term treatment may be optimal for many patients with PDs. However, to provide some more color for clinical guidelines, we must identify and investigate interventions in these treatments and evaluate whether they are in line with the proposed treatment model.

2.17   Treatment fidelity

The lack of adherence and competence scales for documenting treatment integrity undermines the claim of the effectiveness of evidence-based psychotherapies (Perepletchikova, 2007, 2009; Perepletchikova et al., 2007). Wampold and Imel (2015) highlight this by stating that "It is now virtually required that clinical trials of psychotherapy assess and report adherence and competence" (p. 233). Treatment fidelity is defined as the extent to which a treatment is carried out as intended. It includes several interrelated components (e.g. therapy adherence, therapist competence, and therapy differentiation). Nevertheless, adherence alone does not seem to be related to outcome (Webb et al., 2010), must be applied flexibly (Owen & Hilsenroth, 2014), and appears to be a function of the individualities of the patient (Boswell et al., 2013; Imel et al., 2011). Some have argued that therapist effects can be eliminated through proper training and adequate adherence. Results have indicated that the use of a treatment manual and more experienced therapists were associated with small differences between therapists, whereas more inexperienced therapists and no treatment manual were associated with larger therapist effects (Crits-Christoph et al., 1991). Theoretically, from a CF perspective, treatments without any structure, even if the developers have a rationale in mind, will be less effective than treatments that provide the patient a rationale and a plan to overcome their difficulties (i.e., treatment actions consistent with the rationale for treatment; Wampold & Budge, 2012). However, even in investigations that found a relationship between treatment integrity and outcome, the proportion of variance explained by specific techniques was between 10% and 30% (Wampold & Imel, 2015). Thus, the majority of variance was not explained by intended techniques.

2.18 The relationship between adherence, competence, and outcome

Adherence and competence have been proposed to be the best predictors of alliance, which again is the best predictor of outcome (Hatcher, 2010). However, studies on the relationship between adherence, competence, and outcome have been mixed (Barber et al., 1996; Barber et al., 2006; Barber et al., 2007; Beutler, 2004; Castonguay et al., 1996; Crits-Christoph & Connolly, 1999; DeRubeis & Feeley, 1990; Fonagy, 1999; Giesen-Bloo et al., 2006; Høglend et al., 2006; Wampold & Imel, 2015). Some studies support the efficacy of integrative treatments, such as Safran's approach to working with patient–therapist intersubjective experience (Safran & Muran, 2000), which favor therapist tailoring treatment to the patient needs rather than strict adherence to manualized protocols as predictors of overall outcome (Katz et al., 2019; Owen & Hilsenroth, 2014). A recent meta-analysis investigating the correlation between treatment integrity and youth client outcomes concluded that there appeared to be a small correlation between treatment integrity and outcome (Martinez, 2020). An extensive meta-analysis (Webb et al., 2010) reported a small and non-significant aggregate correlation between adherence and outcome (r = .02, 95% CI: -.07–.10). The collective correlation of competence and outcome was also small and non-significant (r = .07, 95% CI: -.07–.20). Boswell et al. (2013) studied the process and outcome of 21 therapists delivering cognitive behavioral therapy (CBT) to 276 patients with panic disorder in a multisite RCT and reported a non-significant overall correlation between adherence and levels of symptoms in the next session (r = .08, 95% CI: -.02, .07). Interestingly, there was a small but significant correlation between competence and subsequent symptoms (r = .15, 95% CI: .05, .25). In general, adherence is not positively linked to outcome, and nor is it negatively linked to outcome (with a few exceptions), indicating that even less effective therapists can be adherent. In a recent pilot study, Esposito et al. (2020) concluded that "adherent interventions should be provided with high quality" (p. 1). However, it remains unclear what such high-quality/competent interventions actually look like, and the authors summarize that "we may claim that clinicians' skills in adapting treatment to the clients and context might be more effective than close adherence to a treatment manual per se" (p. 9). Therefore, it seems that effective therapists adopt their adherence into a real relationship by being "competent in adapting treatment to the clients and context", which seems to be somewhat conceptually related to the working alliance. If such a statement holds any truth, then it seems that Esposito et al. (2020) indicate that adherence necessitates a good working alliance. However, this needs further clarification, and an agreement on such vague concepts may be (almost) impossible and/or undesirable.

Over the past decades, a number of evidence-based treatment approaches for BPD have been developed (Storebø et al., 2020). Nevertheless, interpersonal hostility (Boswell et al., 2013), emotional dysregulation, and relational ambivalence (reactivity) that are often displayed in BPD pathology are factors that may impact adherence and competence (e.g., lead a therapist to become deskilled) in delivering these potions. This has received little empirical attention and may be useful for adapting our theoretical understanding to improving implementation of these treatments. One central concept concerning how skilled therapists navigate sessions on the basis of all available knowledge of the patient, the diagnosis, the treatment, and the relationship, is captured in the term strategic competence (Killingmo et al., 2014). This term and the term embedded alliance are the two central cornerstones in the current dissertation.

2.19   Borderline personality disorder

BPD is a serious, debilitating, and costly psychiatric condition (Choi-Kain et al., 2020; Gunderson et al., 2018b; Leichsenring et al., 2011; Luyten et al., 2020b; Sharp et al., 2020). Studies estimate its prevalence at around 1%–6% of the general adult population and 10%–12% of psychiatric outpatients (Ellison et al., 2018; Torgersen et al., 2001). The disorder is notably common among individuals presenting for medical care in other settings (Zanarini et al., 2004). Individuals with BPD display difficulties with social and occupational functioning and demand a high degree of social assistance (Gunderson et al., 2018b; Zanarini et al., 2009). Traditionally, PDs have been considered difficult to treat due to complicated transference and countertransference reactions with unanticipated issues (e.g., Millon 2004). The presence and increasing severity of PDs have been found to have considerable negative influence on patients' experience of life quality. However, today there exists a remarkable assortment of evidence-based psychotherapies for BPD compared to many other forms of psychopathology, despite persistent therapeutic pessimism (Ellison, 2020). Levy (2008) asks "[w]hat kind of outcome can we expect in the treatment of BPD?" Linehan et al. answer "a life worth living" (Antonsen, 2016), which seems to echo the focus on self-injury and suicide in dialectical behavior therapy (DBT), the most researched treatment for BPD. Work and love appear central in a life worth living, and early on "Freud suggested that mental health depends on the capacity to love and to work" (Daniell, 1985, p. 48).

However, even though all evidence-based treatments for BPD demonstrate symptom remission, far less is known about how to facilitate sustained adaptive work and relationship

functioning. "General functioning may improve, but many treated individuals are still underemployed and lack sufficient stable and meaningful romantic and social relationships at the end of treatment or at follow-up" (Ellison, 2020, p. 424). This is problematic, as we contemplate that a return to normal functioning is considered the consensus standard for treatment efficacy (Lambert & Ogles, 2009). Further, the lack of knowledge about long-term effects of treatment for BPD is particularly troublesome in effective treatments for BPD. "The longest follow-up assessments to date were 5 years after the termination of therapy" (Ellison, 2020, p. 423). The evidence-based treatments come from different theoretical backgrounds and adhere to different assumptions about BPD pathology. Even though these treatments target different symptoms, use different techniques, and are designed with different structures and settings, we know little about which of the evidence-based treatments works best for which patient (Ellison, 2020). As few of the evidence-based treatments report adequate measures for treatment fidelity (the definition of adequate measures for treatment fidelity can be debated), we know little about the causal relationship between treatment and outcome and whether "some therapists win and some do not, independent of the therapeutic method they use" (Nissen-Lie et al., 2010, p. 627) despite serving evidence-based potions for BPD. One study (Kivity et al., 2019) tried to measure conformity to prototypical therapeutic principles and its relationship to change in reflective functioning in three evidence-based treatments for BPD. However, it may be somewhat problematic to consider a study based on the Psychotherapy Q-sort (PQS) (Jones, 1985), a broad-spectrum instrument for characterizing what happens in a therapy session (in terms of CF), a particularly informative adherence study, for reasons that will re-emerge in the discussion.

2.20   The four major evidence-based treatments for borderline personality disorder
Zanarini et al. (2009) indicated four comprehensive psychosocial methods of treatment for BPD. Two of these therapies, MBT and transference-focused psychotherapy (TFP), are considered psychodynamic. The other two, DBT and schema-focused therapy (ST), are cognitive-behavioral. The suggestion that it is the structuring aspect of adherence (e.g., counteracting "entropy"; Karterud et al., 2019) and not adherence to core theoretical ingredients that predicts outcome (Duncan et al., 2010; Wampold & Imel, 2015) seems in line with the observation that completely different theoretical ingredients and interventions in evidence-based treatments for BPD share a common denominator in terms of an overall treatment program tailored to foster a strong working alliance. DBT builds a working alliance through a stable focus on motivation and preventing dropout. Providing therapists with

extensive and detailed manuals, this treatment has a wide arsenal of cognitive, behavioral, and problem-solving interventions, including skills training and in-session coaching, chain analysis, contingency management, and exposure. Such interventions are paired with mindfulness and other practices from Zen Buddhism, with the ultimate aim of decoupling extreme emotional experiences from maladaptive behavioral responses. Another of the four major evidence-based treatments for BPD, TFP, focuses on the disordered mental representations of self and other issues experienced by BPD patients. According to TFP theory, these representations are typically split in all-positive and all-negative parts, which makes it difficult for the patient to regulate their emotions and behaviors effectively, especially in an interpersonal context. Anchored in the present moment, TFP interventions point out conflicting elements of the patient's views of self or others. The therapist also suggests more realistic views and is instructed to take a highly active stance. TFP relies on a strong treatment frame, with a clear delineation of patient and therapist responsibilities; patients are required to pursue meaningful activity (e.g., employment), to maintain sobriety, and to participate fully in treatment sessions.

Psychoeducation may be one crucial element in effective BPD treatments. Zanarini and colleagues (2018) have shown that mere psychoeducation about BPD (i.e., its clinical characteristics, course, etiology, and available treatment options) is better than no-treatment control, at least in some outcomes. ST sees BPD as stemming from a characteristic set of stereotypical schema modes or patterns of thinking, feeling, and behaving with early developmental roots. Lining up with a psychoeducation approach, ST focuses heavily on the provision of a strong, quasi-parental relationship between patient and therapist, something that may facilitate social learning. Bateman et al. (2018) place social learning (epistemic trust) not only at the core of MBT but call it a universal principle for PD treatments (e.g., p. 46), which means there is a need for empirical investigations of the role of epistemic trust in MBT (Paper II). Sharp et al. (2020) argue that the proliferation of MBT lags behind that of more skills-based therapies (e.g., DBT, ST) due to a lack of concrete operationalization of its key components, which is one reason to conduct the current study. Zanarini et al. (2007) argued that DBT and MBT are effective for the treatment of more acute or state-like BPD symptoms, but not for the more temperamental symptoms of BPD.

2.21   The working alliance in effective borderline personality disorder treatments

"Clinicians routinely note the challenges involved in psychotherapy with individuals with BPD, yet little research exists on the therapeutic alliance with this population" (Levy et al., 2010, p. 413). Individuals with PDs often undermine the working alliance (Benjamin & Critchfield, 2010, p. 132). Masterson (1978) suggests that "in psychotherapy with the borderline patient the therapeutic alliance is a goal or objective rather than a precondition" (p. 437). Barber et al. (2010) argue that a "strong therapeutic alliance may be an appropriate therapeutic outcome for certain types of patients (e.g., a patient with BPD or a patient with profound levels of trauma who experiences difficulties trusting or working with others)" (p. 38). Bordin perceived alliance as a vehicle that enables and facilitates specific treatment techniques (Horvath & Greenberg, 1989). Thus, the alliance is embedded within the specific treatment method (Bordin, 1979). The goals and tasks specified appear intimately linked to the nature of the relationship between therapist and patient.

> For example, the kind of bond developed when a therapist presents a patient with a form and asks him to make a daily record of his submissive and assertive acts, and of the circumstances surrounding them, appears quite different from the bond developed when a therapist shares his or her feelings with a patient, in order to provide a model, or to provide feedback on the patient's impact on others. (Bordin, 1979, p. 254)

In treating poorly functioning patients with PDs, the therapeutic stance—being empathetic, attuned, honest, and curious—may facilitate and help maintain a bond between patient and therapist (Bateman & Fonagy, 2016). However, tasks and goals also seem of superior importance, and "all effective treatments share the characteristics of consistency, coherence and continuity, qualities particularly relevant to borderline personality disorder" (Bateman et al., 2018, p. 44). Research has found that improvements in the alliance lead to a reduction in BPD pathology (Levy et al., 2010; Spinhoven et al., 2007). However, within the field of evidence-based treatments for BPD, we have found no investigations of the three facets of the working alliance. Further, within the larger field of psychotherapy research, few have studied the subparts of alliance (Stiles & Goldsmith, 2010), which motivated the present study (Paper II and Paper III).

2.22   Mentalization-based treatment

MBT is a manualized treatment originally designed to treat BPD. It has been found efficient for the treatment of BPD in RCTs and naturalistic studies (Bales et al., 2012; Bateman & Fonagy, 2001, 2009; Kvarstein et al., 2019; Kvarstein et al., 2015; Rossouw & Fonagy, 2012).

Except for a recent study from Denmark by Beck et al. (2020), the two publications by Kvarstein et al. (2015, 2019) remain the only studies of MBT that have included measures on treatment fidelity: The most recent RCT on MBT simply stated "we lacked at the time formal tools to rate therapist competence and fidelity" (Carlyle et al., 2020, p. 7). MBT is a multicomponent treatment combining psychoeducation, MBT group therapy (MBT-G), and individual therapy (MBT-I) and is specifically structured for poorly functioning patients with BPD. The ultimate goal of this psychodynamic treatment is to increase reflective functioning (RF) and move the patient toward greater capacity for secure attachment, thereby enabling effective relationships and better affect and behavior regulation (Ellison, 2020). MBT has been successful in bringing psychodynamic thinking back into the mainstream (Cristea et al., 2017), and "[t]reatment effects are achieved through restoring a balance between the different polarities of mentalizing (automatic versus controlled, self versus other, internal versus external, cognitive versus affective), and by the therapist maintaining a 'mentalizing stance'" (Sharp et al., 2020, p. 2). MBT places emphasis on a therapeutically generated hypothetical change in social communication patterns of increased trust across the spectrum of the client's interpersonal experience (Fonagy & Allison, 2014). MBT has recently placed social learning front and center in understanding mentalizing by introducing the term epistemic trust (Bateman et al., 2018; Bateman & Fonagy, 2016; Bo et al., 2017; Fonagy & Allison, 2014; Fonagy et al., 2015a; Fonagy et al., 2019; Fonagy et al., 2015b; Luyten et al., 2020b; Sharp et al., 2020): "[T]he 'borderline mind', and related severe problems with social communication typically observed in what is commonly referred to as 'personality pathology', may best be understood as a socially triggered outcome based on a learned expectation about the social and interpersonal environment" (Bateman et al., 2018, p. 46). Mentalization refers to how humans make sense of their social world by making inferences about their own and others' mental states, an ability that seems to necessitate epistemic trust. MBT and empirical work suggest that attachment trauma, whether real or perceived, obliterates epistemic trust, shutting down the central "highway" (Fonagy et al., 2015a) for receiving self-relevant information about the world. Sharp et al. (2020) seem to view epistemic trust as the cornerstone for therapeutic change and as essential "in operationalizing the process of rebuilding mentalizing using these observable, behaviourally anchored concepts focusing on creating epistemic trust" (p. 1).

2.23   Mentalization as a core mechanism of change

As we remember from above, mentalization—the ability to mind others' minds, to understand misunderstandings, and to see oneself from the outside and others from the inside—has been proposed as a fundamental CF among psychotherapeutic treatments. It is defined as the ability to understand and interpret, implicitly and explicitly, one's own and others' behavior as expressions of various intentional mental states (e.g., thoughts, feelings, desires). People are very different with respect to mentalizing capabilities. For most of us, mentalizing collapses only sometimes. For others, mentalizing is very difficult most of the time. Poor mentalizing is connected with poor social functioning and psychopathology (Fonagy et al., 2002). The operationalization of mentalization in research is most commonly the (RF; Fonagy et al., 1998) scale. The core assumption in MBT is that an increase in RF will mitigate BPD, and improvements in RF are indeed indicated in effective BPD treatments (De Meulemeester et al., 2018; Levy et al., 2006). The interpersonal process pursuing an open exchange of minds in an attachment relationship with the therapist is assumed to be an effective means to increase mentalization (in borderline patients). Mentalization is thus believed to be facilitated by the quality of the attachment relationship (Fonagy et al., 2002). Thus, the ability to reflect about own's own and others' minds will not develop unless being minded by another human.

2.24   Epistemic trust

Sperber et al. (2010) and later Fonagy (e.g., 2015) and colleagues knowingly borrowed the term epistemic from Aristotle (epistémé; Schwartz, 2011), as "before coining new terms, it is always advisable to look in a dead and learned language to see whether it might not contain such a concept and its appropriate expression" (Kant, 2007, p. 297). However, the merger of Western psychology and Zen Buddhism has proven productive for DBT, and there are probably similar reasons why when investigating proposed universal principles in psychology, such as epistemic trust, there is a long tradition of cross-cultural studies (Passer & Smith, 2004). Therefore, let us first observe that within Indian epistemology, philosophers go into long discussions about apta-vakya ("true knowledge from true sources") and on what grounds they should trust the testimony of the āpta and just how far such trust should extend. The Nyāya school asserts that all forms of valid knowledge are valid only by reason of extrinsic causes, a position known as parataḥprāmāṇyavāda (Hatcher, 1999, p. 64; Picascia, 2019), which roughly translates to "hetero-epistemic theory" (Ram-Prasad, 2013) or "extrinsic validity of cognition" (Shida, 2011). However, we should also observe that Indian epistemology includes (and transcends) Aristotelian logic in its logical tetralemma (Langjord, 2009), such that the Western mind, based on the theory of the four quadrants, will conclude

that epistemic trust not being in the upper-right quadrant (e.g., can be explained by biology), such as Fonagy et al. (2015a) seem to suggest, does not mean that it cannot be part of the upper-left quadrant (interior individual perspective) instead. A somewhat similar logic will also inform us that the fact that mentalizing plays a significant role in evidence-based BPD treatments does not imply that it is the underlying shared mechanism at work in these treatments. Acknowledging this, Bateman et al. (2018) propose epistemic trust for this purpose: "For this we consider it necessary to recognize how individuals 'learn' or fail to learn about themselves and the social world" (p. 45). They further highlight three communication systems central for amending epistemic trust: 1) communication system 1: the teaching and learning of content; 2) communication system 2: the re-emergence of robust mentalizing; and 3) communication system 3: the re-emergence of social learning. The idea that therapy is primarily a learning arena in some form is far from new; for example, Bohart (2000) identifies five "learning opportunities" provided by therapy. Closely related to epistemic trust, psychoeducation, and skills-based strategies, the importance of a pedagogic stance seems implied in evidence-based BPD treatments. Such a pedagogic stance seems somewhat contrary to the "not-knowing stance" championed by a more rigorous MBT. Therefore, at some point one might suspect that the architects of MBT have "thrown the baby out with the bathwater", that is, that their attempt to avoid psychoanalytic interpretations has excluded the focus on the art of transmitting knowledge (typically denoted pedagogy). However, as psychoeducation is already a crucial part of the MBT program (Ditlefsen, 2020) and we have seen that psychoeducation in itself (e.g., Zanarini et al., 2018; Zanarini & Frankenburg, 2008) shows good effect for BPD patients, there are indications that a pedagogic stance may be central for MBT.

2.25 Epistemic trust: A "royal road" to understanding psychotherapy?

In an attempt to reintroduce the baby—hopefully without too much bathwater—a "pedagogic stance" has recently been advanced as an MBT intervention in the Scandinavian manual for MBT (Karterud et al., 2020). Transmission of knowledge means that the overarching strategy should be in accordance with the patients' level of insight or (lack of) knowledge (Goldfried, 2008). This is probably the most established principle of pedagogy; for example, according to Aristotle, "[a]ll teaching is from things previously known" (Schwartz, 2011, p. 119). However, "what is grasped by epistēmē (epistētoñ) is what is demonstrated, and since there have to be first principles of demonstration, there is no epistēmē of the principles of knowledge. That is to say, principles of demonstration cannot themselves be demonstrated"

(Schwartz, 2011, p. 119). Or, as Alan Watts eloquently put it, "the knower is never an object of its own knowledge" (1999, p. 69), to which he added "because fire does not burn fire" (Watts, 2004, session 6).

Therefore, opening the channel for epistemic trust or fostering "the miracle of understanding" (Gadamer et al., 2004, p. 309) is a sophisticated art, and there are good reasons why pedagogic interventions of poor quality should be avoided. First of all, therapists need to avoid preaching what they themselves need to hear or trying to convince the patient of some of their own opinions. This would be considered toxic in terms of MBT quality and would most likely be awarded a rating of 1–2 on a scale of 1–7 (Karterud & Bateman, 2010; Karterud et al., 2020; Karterud et al., 2013). Therefore, introducing the pedagogic stance in the MBT framework comes with substantial risk, as do all ambitious (MBT) interventions, such as the use of countertransference or challenging unwarranted beliefs (e.g., Piper et al., 1991). At this point, it is perhaps good to ask what MBT actually is. Disregarding the problem of the chicken and the egg for a second, one might, for instance, ask if MBT is what is defined by the manuals and thus somehow abides among the platonic forms publicized by the architects of the treatment (theoretical) and/or whether is it better described by what MBT therapists are actually demonstrating (empirical). If we lean slightly towards the upper-right quadrant in this question, it becomes of interest to see what we can learn from 327 MBT individual sessions rated by the Quality Lab for Psychotherapy at Oslo University Hospital (Table 1). Surprisingly, Folmo et al. (2021b) and Karterud et al. (2020) found that the most prevalent intervention used by Nordic MBT therapists is "validating understanding" (Item 16; Karterud et al., 2013). In fact, this intervention accounted for 32% of the identified interventions. This simple finding is remarkable, given that this item even surpasses "exploration, curiosity and a not-knowing stance", which should be the hallmark of MBT. Shifting our focus to the upper-left quadrant, our best explanation of this is that therapists use this item to be (indirectly) pedagogic, often in a concealed way, because interpretations are to be avoided according to the manual (Karterud & Bateman, 2010). Therefore, despite MBT being operationalized as a "not-knowing" therapeutic approach, pedagogic interventions (just like in the other three evidence-based treatments for BPD) pervade the actual therapies performed and also seem to color all other intervention types to various degrees (ibid). Employing interpretative phenomenological analysis (IPA) to nine MBT-G and 24 MBT-I sessions, Folmo et al. (2021b) concluded that MBT seemed to mainly address communication systems 2 and 3 (Bateman et al., 2018), while the more skills-based treatment for BPD may

involves system 1 (learning of content). Folmo et al. (2021a) also identified what seemed like nine prototypical versions of interventions targeting impaired epistemic trust (including missed opportunities); see Table 2. The possible existence of such discrete categories for pedagogic interventions (targeting specific domains of impairment in epistemic trust) signals that it may be possible to identify a limited number of overarching strategies for pedagogic stance, even within a psychodynamic and understanding-driven approach as MBT. Folmo et al. (2021a) indicated that pedagogic interventions strengthened the alliance and epistemic trust in MBT, and "pedagogic stance" is now proposed as an intervention in MBT (Karterud et al., 2020).

2.26   Pillars of alliance in mentalization-based treatment

In light of the above discussion, it becomes especially interesting to ask what it is that connects the specific ingredients in the MBT potion with the CFs (i.e. alliance). As alliance in other psychotherapies has been referred to as the "quintessential integrative variable" (Wolfe & Goldfried, 1988, p. 449), it is of interest to us how one can establish a stable expression of the relationship (alliance) with patients whose primary pathology is substantial problems with making and maintaining stable relationships. The very low dropout rate reported in MBT (Kvarstein et al., 2015) indicates that patients experience a good working alliance with the treatment, structure, and most likely the therapists. How does the therapist co-create a working alliance with someone with relational pathology and disbelief in others' knowledge, that is, disturbed attachment and low epistemological trust (Fonagy & Allison, 2014)? And what does it look like when this is rated highly or poorly according to the adherence and competence scale for individual MBT (Karterud et al., 2013)? BPD patients are likely to be in for a long and tough ride in therapy, and one might reason that agreeing on the goals and tasks and experiencing a safe personal bond is paramount for these patients. BPD patients display schematic, rigid, and sometimes extreme views (Gunderson et al., 2018b), which an effective treatment needs to address and challenge. Therefore, to prepare and deliver a potent "magic potion" to this population, we need a larger system (the MBT program), highly trained and competent therapists, and a clearly defined recipe (manuals). MBT programs consist of four structural pillars built to establish a "strong alliance": 1) psychoeducation, which is an important tool in agreeing on goals and tasks because it explains central features of BPD, mentalizing, affect, attachment, and the treatment program (Karterud, 2011, 2019); 2) an individual dynamic MBT case formulation (Karterud & Kongerslev, 2019a); 3) individual mentalization-based psychotherapy (Karterud & Bateman, 2010; Karterud et al., 2020); and 4)

MBT-G (Karterud, 2015). All patients are required to participate in 12 sessions of psychoeducation when enrolling in the MBT program, and the overall focus is on integrating all aspects of the treatment. For example, the individual therapist encourages the patient to attach to the psychodynamic group and vice versa. The manual for individual MBT (Karterud & Bateman, 2010) states that "If a patient drops out of one of the components, then the other components are automatically terminated" (p. 42). Without such clear practice, the therapist will easily fall victim to borderless borderline patients whose personality is often specialized in pushing, pulling, and forcing others into their own scripts/schemas/patterns. Therefore, counteracting these forces, the alliance-fostering ingredients in the MBT potion are arguably stronger (more caring, strict, pushy, normative, and committed) than in many other psychotherapy orientations and are based on goals that the patient deeply commits to. The task the patient is embarking on (radical change in personality) is also considerably more difficult than in therapies with less disturbed patients.

2.27   Mentalization-based treatment or plain old therapy?

MBT has been called both "plain old therapy" (POT) and (purified) old wine in new bottles (Allen, 2012). The novelty of MBT is that it keeps a steady focus on mentalization, which is the basic ability to understand relations, inner processes, and guesstimate others' mental content (the specific ingredient in the "magic potion"). However, an increase in mentalizing ability would most likely be the hallmark of all effective therapy with a focus on relationships, feelings, and self-understanding. The purification, then, is the clear focus on mentalizing, mental states (minding minds), the absence of interpretations (performing the mentalizing for the patient), defocusing on insight or historical content, and an unwavering spotlight on the 17 colors in the spectrum of building a strong alliance. MBT is a manualization of a non-technique-based psychotherapy (Perepletchikova et al., 2007; interventions are driven by understanding), in which the relationship with the therapist and interactional processes play a central role. Therefore, in addition to the clear focus on the 17 constituents, the manual also states that "The therapist must offer himself/herself as a possible attachment figure, thereby becoming emotionally involved in the patient's life. The therapist must 'care'" (Karterud & Bateman, 2010, p. 43). In individual MBT, this strong alliance is differentiated into 17 core activities the individual therapist must employ to facilitate the long journey from having a personality disorder to manifesting "a life worth living". Simply stated, one could say that the "strong alliance" (patient and therapist agreeing that the goal is to improve mentalizing and reduce BPD traits) can be broken down, as light through a prism,

into these 17 defining items in individual MBT (Karterud et al., 2013) and 19 in MBT-G. These ingredients are different ways to obtain an intact and efficient alliance in MBT (e.g., challenging, exploring in a not-knowing way, validating, acknowledging, displaying genuine interest, exploring the therapy relationship, stopping pretend mode, stopping psychic equivalence, being open and transparent about one's own mind). As these 17 and 19 elements comprise a totality, address different cornerstones for change, and prevent possible escape routes for the patient (e.g., "pretend mode"; pretending to be normal), it is important that all are applied (in a way tailored for the patient). This means one can have a high adherence score on individual MBT even though not employing all 17 items in one session, but a certain sign of poor MBT would be neglecting one of these areas if it were indicated in the session. One central aim of the current study was to investigate if observers can agree on the 19 ingredients in MBT-G (Paper I).

2.28   The problem of manualizing mentalization-based treatment

Despite MBT supposedly being based on "the most fundamental common factor among psychotherapeutic treatment", it has been criticized for being too abstract and relying too heavily on expert supervisors who can translate dense psychodynamic theory into practice. Hutsebaut et al. (2012) reported in their implementation study that MBT-trained therapists felt insufficiently prepared to apply their new knowledge and skills in everyday practice. MBTs include some skills-based learning and general strategies, such as increasing mentalizing flexibility by regulating emotional activation and being transparent about one's own mind. MBT manuals suggest curiosity, a high level of genuine care, intellectual humility, low rigidity, and high tolerance for transferences as core facets of a mentalizing stance, "but granular-level, behaviorally anchored guidance is not provided on how to achieve these" (Sharp et al., 2020, p. 3). As a consequence of this lack of specificity, MBTs may be difficult for novice therapists to learn. Sharp et al. (2020) suggest that "[c]oncrete protocols may be needed to reduce therapists' uncertainty and anxiety" (p. 3). However, many specific examples of behavior in manuals for psychodynamic therapies could make the treatment rigid, and slavish adherence to treatment protocols has been suggested to result in deterioration of the therapeutic relationship (Henry et al., 1993). Manualization may also become too rigid by generalizing principles that are valid for the vast majority of patients but are not necessarily applicable to the specific patient. Therefore, the core of MBT, like most psychodynamic therapies, is non-directive and non-instructional; a "one size fits all" approach would be somewhat in contrast to the core of co-creating a safe learning environment with the

unique patient (working alliance). All three core components of MBT are manualized (Bateman & Fonagy, 2016; Karterud, 2015, 2019; Karterud & Bateman, 2010; Karterud et al., 2020), but it is hard to manualize how to create and maintain a mentalizing therapy culture amongst therapists. The treatment presupposes a well-functioning team and video supervision services that manage to integrate the different components and different therapists involved (Bateman & Fonagy, 2016, pp. 155–156), and organizational disruptions may affect the outcome of MBT (Bales et al., 2017a; Bales et al., 2017b). Therefore, successful MBT demands a system that supports these challenges where the therapist must redefine and rediscover her/his role as therapist and attachment figure with every patient.

2.29   The patient's contribution to competence

It has generally been assumed that adherence and competence are therapist characteristics (Baldwin & Imel, 2013). Wampold and Imel (2015) summarize their understanding of the Boswell et al. (2013) study thus: "Although not quite statistically significant, it does show that it is the patient's contribution to competence ratings that is related to outcome rather than the therapists' competence relative to other therapists" (p. 238). MBT is a dynamic psychotherapy, and the manual is based on and driven by understanding. Therefore, the competent MBT therapist must be interpersonally skilled, able to work collaboratively with a range of patients, express empathy (therapist empathy is a core specific ingredient in the treatment), and effectively engage the client in the treatment actions. This indicates a need to investigate the "role of responsiveness in treatment adherence and competence with particular patients (e.g., when and why a therapist 'goes off track' with a given patient), including the immediate and direct impact of patient characteristics on therapist behaviour and decision making" (Boswell et al., 2013, p. 453). In MBT, it would seem plausible that individuals with different pre-treatment levels of mentalizing capacity may differ in their ability to engage in psychotherapy (Katznelson, 2014), and different capacities for mentalization need different kinds of therapeutic approaches (Antonsen, 2016).

2.30   Adherence and competence/quality for mentalization-based treatment

In an exemplary RCT, "based on video recordings of the therapy sessions, independent observers assess how closely the therapists adhere to the treatment manuals (adherence), and how competent they are (specific therapeutic competence)" (Jørgensen, 2019, p. 53). A recent review of the evidence-based status of MBT (Malda-Castillo et al., 2019) found that fidelity to treatment was poorly reported in almost half of the studies (47%). Importantly, there is

currently no consensus as to what counts as good fidelity measures, which is why the current dissertation reports different numbers here. Based on the Norwegian manual for individual MBT (MBT-I; Karterud & Bateman, 2010), an adherence and competence scale (MBT-I-ACS) was developed (Karterud et al., 2013). It provides possibilities for the documentation of model fidelity in treatment studies (e.g., Kvarstein et al., 2019; Kvarstein et al., 2015). It has also been used for in-session studies of therapists' interventions and their relationship to outcome (Möller et al., 2017). The Quality Lab for Psychotherapy has implemented the MBT-I-ACS as part of its quality control system, and the scale has also been employed for educational purposes. Recently, MBT-G has been manualized (Karterud, 2012, 2015) and provided with practical guidelines (Bateman & Fonagy, 2016). Karterud's manual (2015) follows the recommendations of Luborsky and Barber (1993) and includes (1) a theoretical rationale; (2) presentation of the main principles underlying the therapeutic techniques; (3) concrete examples of all techniques being described; and (4) scales and instruments that can assess the skills of the therapists for this particular treatment model. The manual also contains a 19-item adherence and quality scale for MBT-G (Karterud, 2015). So far, the scale has been used in process studies and to explore similarities and differences between psychodynamic group therapy and MBT-G (Beck et al., 2020; Kalleklev & Karterud, 2018; Karterud, 2018).

2.31   Measuring mentalization-based treatment: The current dissertation

To investigate the presented themes further, this thesis will examine one of the evidence-based treatments for BPD, primarily because it is a manualized, dynamic psychotherapy where the curative ingredient, mentalizing, has been theorized to be a CF among psychotherapeutic treatments. Therefore, as the specific curative ingredient in MBT is considered a CF, MBT seems an ideal candidate to explore the relationship between specific factors and CFs (as we can view the same ingredients through both lenses), an overarching aim of the present study. Further, even though evidence-based treatments for BPD may seem like a limited field to investigate, such a study is worth our attention for two main reasons, as follows. 1) As laid out above, evidence-based treatments for BPD are an exception to the rule when it comes to demonstrating superior effect compared with treatments as usual (Ellison, 2020). A previous dissertation (from Oslo University Hospital, Ullevål) within the field of psychotherapy for PD states that future research should emphasize treatment in accordance with *what* works for *whom* (Antonsen, 2016). Therefore, it is of interest to study how MBT therapists tailor their specific technique to each individual patient. 2) Everybody has a personality, a characteristic manner of thinking, feeling, behaving, and relating to others

(Mischel & Shoda, 1995; Westen, 1995), and even though BPD patients are characterized by insecure attachment, unstable emotions, poor mentalizing, and impaired epistemic trust (Bo et al., 2017), it seems likely that all patients can experience such phenomena at times. It is therefore possible to argue that the results of measuring MBT could be translated to and be of interest for the broader society for psychotherapy research and for practicing psychotherapists.

2.32   Measuring mentalization-based treatment: What do we know?

Today, psychotherapy is generally seen as the most successful treatment approach, and several evidence-based psychotherapies for BPD exist (Cristea et al., 2017). However, this was not always the case, despite the existence of DBT (Linehan, 1993), it was a major breakthrough when Bateman and Fonagy (1999, 2001) reported good effects of their "psychoanalytically oriented partial hospitalization" model for BPD patients. This "psychoanalytically oriented" treatment was later relabeled MBT. The impressive results could not necessarily be attributed to the treatment model, as 18 months of partial hospitalization could imply countless reasons for clinical change. The process of defining and manualizing this treatment resulted in an intensive treatment model composed of (1) a psychoeducational group (2–3 months), (2) individual therapy once a week, and (3) mentalization-based group therapy once a week for about 18 months (Bateman & Fonagy, 2006). Following this program, MBT demonstrated superior effects compared with "structured clinical management" in a RCT with 134 BPD patients (Bateman & Fonagy, 2009). However, the implementation of MBT has had various degrees of success in Northern Europe. Jørgensen et al. (2013, 2014) found general improvement among included patients but no superiority. However, 48% of the patients who received 2 years of combined MBT treatment met criteria for functional remission (GAF-F>60) at 1.5-year follow-up, compared with 19% of the patients who received supportive group therapy (Jørgensen et al., 2014). Even though this difference in functional remission only approached statistical significance, it may be clinically significant. Over a 8-year follow-up period, "participants treated with MBT showed better functional outcomes in terms of being more likely to be engaged in purposeful activity and reporting less use of professional support services and social care interventions" (Bateman et al., 2020, p. 1), and it may be that "MBT provides long-term improved outcomes, which are often not investigated in other psychotherapies" (Volkert et al., 2019, p. 25). In the Netherlands, it was discovered that positive effects of MBT treatment depend on the structure of the entire program (Bales et al., 2017a; Bales et al., 2017b; Hutsebaut et al., 2012).

Consequently, a quality manual for the implementation of MBT with a focus on how to structure the organization was developed (Bateman et al., 2012). In treatment and process studies on PDs, some commonly used outcome measures are: (1) severity of PD, typically indicated by the number of fulfilled SCID-II criteria, or specific BPD measures, such as the Zanarini Rating Scale for Borderline Personality Disorder (Zanarini, 2003); (2) interpersonal dysfunction measured by the Circumplex of Interpersonal Problems (Alden et al., 1990; Boudreaux et al., 2018; CIP; Pedersen, 2002); (3) social functioning often measured by the Global Assessment of Functioning (GAF; Pedersen et al., 2018), number of months or years in work, studies, or measures from the Work and Social Adjustment Scale (WSAS; Pedersen et al., 2017b); (4) symptoms measured by Symptom Checklist-90-Revised (SCL-90-R; Derogatis, 2017; Derogatis & Unger, 2010) or Brief Symptom Inventory 18 (BSI-18; Derogatis, 2000); and sometimes also (5) attachment style as indicated by the Experiences in Close Relationships (ECR; Fraley et al., 2000). Effect sizes around 0.7–0.8 are considered substantial in psychotherapy research and in an MBT study by Kvarstein et al. (2015) effect sizes for the measures CIP, BSI and GAF were all large (>1.0). This was also the first study that reported measures for treatment fidelity. We know little about adherence and competence in other effect studies of MBT (Bateman & Fonagy, 1999, 2001, 2009; Jørgensen et al., 2013; Laurenssen et al., 2018). The Quality Lab for Psychotherapy rated MBT-I sessions in the study by Philips et al. (2018), in which the manual was not followed adequately (mean MBT-I-ACS scores were below 4). Predating the introduction of evidence-based treatments for BPD, the typical dropout rate was around 40%–60% (Karterud et al., 2020). Kvarstein et al. (2015) found a dropout rate of 5% in the first 6 months, a finding that highlights the importance of the alliance in treatments for BPD. MBT, in contrast to treatment as usual, has also been found to maintain treatment effects with increased PD severity (Bateman & Fonagy, 2013; Kvarstein et al., 2019). DBT is typically considered the evidence-based treatment for BPD with most empirical support. However, if one disregards the mere number of RCTs and instead focus on effect, there are only marginal differences between DBT and MBT (Cristea et al., 2017; Oud et al., 2018; Storebø et al., 2020). This is also the case in terms of self-harming, which is more explicitly targeted in the DBT manuals. In Norway, DBT has reported stronger effects (compared to a control group receiving enhanced usual care) for self-harm amongst youth (Mehlum et al., 2014). Further, despite MBT being an effective treatment for BPD (Storebø et al., 2020), few have investigated the core theoretical assumption that an increased ability to mentalize (reflective functioning; RF; Vogt & Norman, 2019) is indeed the mechanism of change in MBT: "Fewer studies have investigated

the purported role of changes in RF in relation to clinical outcome in treatments focusing on this capacity" (Chiesa et al., 2020, p. 1), and the results have been mixed (Chiesa et al., 2020; Goodman, 2013; Karlsson & Kermott, 2006; Katznelson, 2014; Rizq & Target, 2010; Rudden et al., 2006; Taubner et al., 2011; Taubner et al., 2013; Tmej et al., 2018; Tmej et al., 2020; Trowell et al., 2008). In an RCT by Rossouw and Fonagy (2012) for youth (self-harm), the effect of MBT was mediated by an increase in RF. Möller et al. (2017) reported that adherence to MBT principles and competence in the performance of MBT predicted enhanced mentalizing from patients during a session. In terms of RF as an outcome measure, some studies have associated outcome of BPD treatment with improvement in RF (De Meulemeester et al., 2018; Fischer-Kern et al., 2015; Goldstein, 2015; Levy et al., 2006) In a case study of a BPD patient, a significant increase in RF was reported after three years of mentalization-based treatment (Gullestad & Wilberg, 2011), and RF has also been found to increase in MBT-G (Kalleklev & Karterud, 2018). Vermote et al. (2010) reported that inpatients with personality disorders did not change their RF after treatment, despite significant symptom reduction, while RF measured in the first month of treatment predicted changes in two outcome measures of personality functioning post therapy in a study by Boldrini et al. (2018).

2.33   Measuring mentalization-based treatment: Four missing pieces in the jigsaw

In the above presentation of what we already "know" about measuring MBT, we have seen that this evidence-based treatment for BPD is well established within the EST approach. Therefore, one overall aim of the current dissertation is to bridge the CF and EST approaches in our study of measuring MBT. All three components of MBT are manualized, and at the start of the PhD period reliable fidelity measures for individual MBT (MBT-I) (Karterud et al., 2013) were established but not for MBT-G. Karterud et al. (2013) achieved G-coefficients for adherence and competence by seven raters of .84 and .88, respectively, which is very high. If the adherence and quality scale for MBT-G (MBT-G-AQS) could achieve similar reliability, it would allow this instrument to document treatment fidelity for the overall MBT program. Therefore, Paper I investigated the reliability of the adherence and quality scale for MBT-G.

As we have seen, MBT is an EST for BPD patients. However, we know little about what good and poor MBT actually looks like (Paper II), and we know little about how the different components of alliance develop in MBT (Paper III). Paper II applies qualitative methods

(Wilber's upper-left quadrant) to this topic, as recommended (e.g., Elliott, 2010; Kazdin, 2009). Paper III is a quantitative study that employed linear mixed models (LMMs) to investigate what kind of development in the three facets of the working alliance characterized MBT with good and poor outcome.

The three papers in this study aimed to investigate the following four major missing pieces in the jigsaw of measuring MBT:

1) Measurement of adherence and competence/quality is necessary in order to evaluate the effect of any specific treatment, e.g., whether it is reasonable to infer that the treatment in question is MBT? MBT combines individual and group formats, but has lacked methods for assessing adherence and competence of group therapist interventions. In a quantitative study, Paper I examined if reliable treatment fidelity could be achieved for the group component.

2) If reliable measures of treatment fidelity exist (Paper I) would it be possible to explain the differences between highly rated and low-rated MBT sessions without simply applying circular logic in terms of stating that good MBT is characterized by many high-quality MBT interventions according to the manual (i.e., adherent and competent MBT), and vice versa? Paper II investigated whether MBT can be interpreted by primarily employing CF terminology. For example, can therapeutic alliance explain the differences between highly rated and poorly rated MBT?

3) As the working alliance is the major mechanism of change identified in psychotherapy research in general, will aspects of alliance also be associated with outcome variation in MBT specialized for patients with BPD? In a large sampled, quantitative longitudinal investigation, Paper III explored the development of working alliance in MBT with different outcomes.

4) This dissertation has the overall aim of integrating MBT with(in) the larger field of psychotherapy research, where alliance and therapist effects are considered amongst the primary mechanisms of change. The crucial question becomes whether CFs can account for much of what happens in MBT.

2.34 Paper I

Few group psychotherapy studies focus on therapists' interventions, assessment of group therapists' intervention competence and quality is complex as the process also includes member to member interactions and the group as whole, and instruments that can measure group psychotherapy treatment fidelity are scarce. The group component is considered the clinical backbone of MBT. The aim of the study was to evaluate the reliability of the MBT-G-AQS developed by Karterud (2015), which is a 19-item scale to measure adherence and quality in MBT-G. This quantitative study addressed the following two research questions.

1) Can trained raters obtain adequate interrater reliability for ratings of overall adherence and quality and the different items (i.e., can the MBT-G-AQS be reliably rated)?

2) How many raters are needed to obtain adequate reliability for the overall scale and each of the 19 items? (With how few raters can we achieve adequate reliability?)

2.35 Paper II

Given the current state of psychotherapy research where talented therapists evidently produce a stronger alliance without researchers being able to identify what they do (Fonagy, 2010; Lemma et al., 2011), this study aimed to investigate how highly rated and low-rated MBT therapists tailor their treatment to patients. The Quality Lab for Psychotherapy measured 108 individual sessions with the newly developed adherence and competence scale for individual MBT (Karterud et al., 2013). In a qualitative design, our question was what characterized selected sessions with high and low ratings. Avoiding the circular logic of answering the question in terms of the MBT framework, we investigated whether observed differences could be explained in terms of CFs.

2.36 Paper III

Despite attachment issues in the target population, there is little research on the relationship between alliance and outcome in the study of PDs. The effects of MBT are well documented, with low dropout rates and substantial improvement in terms of self-destructiveness, and symptom relief. Central relational problems among patients with BPD are associated with hypersensitivity, insecure attachment, and lack of epistemic trust, aspects particularly challenged in the therapy setting. The overriding aim of this quantitative study was to investigate how aspects of therapeutic alliance (goals, tasks, and bonds) developed over time

in MBT for patients with BPD. The study primarily aimed to investigate alliance processes in therapies with different clinical outcomes and secondarily to explore variation associated with different patient characteristics.

3   Methods

3.1 Paper I

This paper is a quantitative study of the reliability of therapy raters.

3.1.1   Participants

Patients

Approximately 85% of the patients were females with a primary diagnosis of BPD.

Therapists

Eight psychodynamic group sessions (PDGs) and eight MBT-G sessions were included. Fourteen therapists (mean age = 53, SD = 10.4 years, range 27–69 years; 57% females), all regular employees at the Department of Personality Psychiatry, Oslo University Hospital, volunteered to participate in the study as part of their ordinary workload. Twelve were experienced clinicians and group analysts (certified by the Institute of Group Analysis in Oslo). By profession, there were five psychiatrists, one psychiatric resident, two clinical psychologists, one social worker, one student of psychology, one physiotherapist, and three psychiatric nurses. All therapists, except the student and the psychiatric resident, had been trained in MBT locally when the department changed from a group-oriented day hospital program to an MBT program in 2008. Each group had two therapists (a pair of therapists). Average supervision hours every week per therapist in the study numbered four.

Raters and rater training

The ratings were made by five raters who were all trained MBT therapists. By profession, there were four psychologists and one professor of psychiatry. Thorough theoretical and practical training in the MBT-G-AQS was performed by the author of several MBT manuals (Karterud, 2011, 2012, 2015, 2019; Karterud & Bateman, 2010; Karterud et al., 2020). The training included rating two verbatim transcripts of MBT groups and a subsequent discussion of the ratings. Four of the five raters had assessed at least 30 (range 30–91) sessions with the MBT-I-AQS as part of The Quality Lab for Psychotherapy.

Treatment

The patients in both treatment modules (PDG and MBT-G) had been in the program for various lengths of time (e.g., 0–36 months). All groups were slow open, accepting new members at approximately the same pace as patients completed the program. 1) MBT-G: Participants in the MBT program were offered weekly MBT-G for a maximum of 3 years. 2) PDG: The treatment in the Ullevål Personality Project (UPP) was a long-term combination program comprising short-term day-hospital treatment followed by outpatient combined group and individual psychodynamic psychotherapy. A detailed description of the design and treatment setting of the UPP clinical trial is presented elsewhere (Arnevik et al., 2009b).

3.1.2 Selection of tapes

MBT-G sessions were recorded, while PDG sessions were selected from a pool of recordings. Eight sessions were selected by purposeful sampling. To minimize variance due to therapists' general competence, we included two therapists who performed both PDG and MBT-G. Regarding the selection of MBT-G tapes, the four pairs of therapists provided two videotaped recordings each. They were assigned to record two subsequent sessions over a span of 3 weeks. Regarding the selection of PDG tapes, the first step was to identify therapists who had comparable formal education to the four pairs of therapists recording MBT-G. Recordings were then randomly selected within the specified groups (pairs of therapists).

3.1.3 Scale development

The MBT-G-AQS (Karterud, 2012, 2015) is a 19-item scale developed with the purpose of measuring adherence and competence in MBT-G. The scale was based on the previously developed and reliability tested MBT-I-ACS (Karterud et al., 2013) and the three higher-order domains for group therapy defined by Chapman et al. (2010). The items in the MBT-G-AQS are shown in Table 1. The nine group-specific items in the MBT-G-AQS (different from the items in the MBT-I-ACS) were identified and operationally defined based on two criteria: 1) that they should reflect significant motives for group psychotherapy interventions in a more general sense, and 2) that they reflect treatment needs according to the theory and practice of MBT. A large item pool was assembled and critically reviewed. Those items that survived the scrutiny were further defined and debated. A preliminary version of the MBT-G-AQS scale was evaluated by studying video recordings of group sessions from Norway, Sweden, Denmark, and the UK (Karterud, 2012, 2015). Subsequently, a manual was created that described 1) the essence and general principles of MBT-G and 2) the essence of each item,

with detailed indicators for quality ratings and examples of the intervention to facilitate adherence ratings (ibid.).

Content validation

The MBT-G-AQS was developed before the current investigation was initiated. For content validation and to avoid local/national idiosyncrasies, the developmental work with the manual (item selection, item definitions, quality descriptors, rating procedures, etc.) was performed in a dialogue with clinicians and researchers from Sweden (Stockholm) and Denmark (Copenhagen, Roskilde, and Aarhus) (Karterud, 2012, 2015).

3.1.4 Assessment

All therapy sessions were videotaped and assessed by five independent raters. All video sessions were rated by all raters in their entirety. The adherence and competence ratings did not have missing data, as the raters scored all items for all videos assessed. To prevent the raters from drifting during the evaluation process, ratings of every video were compared and the differences were discussed, which is consistent with previous research studies (von Consbruch et al., 2012; Weck et al., 2011). Ratings were not changed after this comparison. Thus, the ratings were assessed independently.

Ratings of adherence and competence

All items in the MBT-G-AQS were rated on an 8-point Likert scale from 0–7 for competence (quality) where "0" is "Not applicable (the intervention was not observed)", "1" is very poor, and "7" is "excellent". Fourteen items were rated on an 8-point Likert scale from 0–7 for adherence. The range is from "not at all" (score 0) to "extensively" (score 7). All raters also assessed global adherence (one item) and global competence (one item). The manual for MBT-G (Karterud, 2015) contains detailed descriptions of the items and scoring procedures. All items are shown in Appendix 1 and described by their quality rating of 4 ("good enough").

Adherence primarily relates to the frequency and extensiveness of prescribed MBT interventions. Frequency means the number of times the therapist carries out an intervention, and extensiveness means the time and attention the therapist gives the intervention. Five of the items ("care for the group and its members", "managing authority", "engagement, interest, and warmth", "regulating emotional arousal", and "handling pretend mode") are not assessed

for frequency. If the therapist fails to deliver clearly indicated interventions, the item can be rated low on quality (e.g., 2) even where there is no (or low) occurrence (adherence/frequency score = 0, 1, or 2). After each item is assessed, the rater decides on an overall score for the specific therapy session for both adherence and competence. This global assessment is not made on the basis of an arithmetic average of the 19 items but according to an overall clinical judgement. The rater's starting point is at "4", indicating the therapist is "good enough" (adequate/average MBT therapist). The rater adjusts their scores in relation to deviations in a positive or a negative direction from this starting point. An overall score of 4 is defined as an adequate performance both in terms of adherence and competence. A low rating means that the therapist did things other than those prescribed in MBT, that they performed interventions in an inflexible or clumsy way, or that they did not follow up interventions adequately. The raters evaluated the conjoint efforts of the observed therapists (not each independently). It is possible for one therapist to be much quieter than the other and achieve high ratings. However, if considered an unintended intervention of being silent (e.g., a "missed opportunity"), then such an imbalance or lack of communication with the co-therapist (Item 9), this would typically account for a reduced quality rating. In fact, in well-cultivated MBT groups, the therapists can remain fairly silent for a long while because a mentalizing discourse has been ignited in the group and/or because the structure of MBT allows for the group to stick to one topic. Therefore, the high frequency of interventions does not always imply high quality, and vice versa. To emphasize that it is the therapists' activity being judged, the manual instructs that qualifying statements for the items should be of the format "to what degree did therapists X and Y do…?" with respect to adherence and of the format "the therapists' interventions were…." or "the therapists did….", etc. with respect to quality. The items are defined using specific clinical examples based around observable behavior whenever possible; for example, "The therapists invite the other group members, implicitly or explicitly, to clarify relevant events and engage members to participate in a collective exploration of the mental states involved therein" (Item 4; see Appendix 1).

3.1.5 Data analysis

Generalizability theory

A much-welcomed approach to the examination of psychotherapy sessions is generalizability theory (G-theory; Cronbach et al., 1963; Shavelson & Webb, 1991), which is suitable to investigate observational ratings of complex phenomena. The data dictates the method, and

when the measurement design contains multiple sources of variance, G-theory is an appropriate approach to disentangle and estimate these sources of variance. G-theory addresses the adequacy with which one can generalize from a sample of observations to a universe of observations from which the sample was randomly drawn. This issue is particularly relevant for ratings of the psychotherapy process because multiple sources of error variance are common, such as within the variation due to the patient, the session, the group, the therapist, the rater, or other potential factors. In this study design where the observed score is compounded by three or more sources of variance, intraclass correlation is not an appropriate method to estimate the level of reliability. By incorporating multiple sources (facets) of error into reliability coefficients, reliability estimates calculated using G-theory are likely to be more accurate, as some contributions to errors of measurement (e.g., occasions, raters or items) can be assessed (Shavelson & Webb, 1991). G-theory provides estimates of the variability contributed by each source of error and of the interactions among sources of error (Shavelson & Webb, 1991). Although G-theory is concerned with variance components, it produces a summary coefficient, the G-coefficient, which is roughly analogous to a reliability coefficient (e.g., intraclass correlations). However, the G-coefficient is based on the researcher's decision to treat facets as random or fixed, thereby defining the universe to which the researcher wants to generalize. Within the design of G-theory, several variance components can be disentangled in just one analysis (Shavelson & Webb, 1991). Further, a generalizability study (G-study) makes it possible to disentangle the component variations and estimate the reliability for a decreasing number of raters. The question in G-theory is the degree to which observed scores allow for generalizations about a person's behavior in a defined universe of situations. G-theory provides G-coefficients reflecting the variability contributed by each source of error and of the interactions among sources of error (Shavelson et al., 1989). Based on the sample data, the relative impact of different sources of variation is estimated by a G-study (Shavelson et al., 1989), from which generalizability coefficients are computed. The G-coefficient (ρ^2) indexes the proportion of total variability in scores that is due to "universe scores" ($\rho^2 = \frac{\sigma^2(\tau)}{\sigma^2(\sigma^2(\tau)+\sigma^2(\delta)\tau)}$), where $\sigma^2(\tau)$ is the variance of the true score, and $\sigma^2(\delta)$ is the variance of the various error components. G-coefficients above .7 are generally considered sufficient for interpersonal and observationally coded constructs (Wasserman et al., 2009). A low G-coefficient is due to a significant amount of error in measurement or to minimal variation across individuals, the measurement procedure, and the universe of generalization (Hagtvet, 1997). The second coefficient in G-theory is called the

dependability coefficient, denoted as D, and can be interpreted as the generalizability coefficient for absolute decisions (Shavelson & Webb, 1991). Therefore, based on the obtained G-study components, the generalizability framework offers a subsequent study called D-study or optimization study. With the D-study it is possible to estimate the reliability of scores based on, for example, four, two, or only one rater(s). This also allows for the more efficient training of raters, as the G-study predicts what the reliability would be if any rater were excluded from the study. If the G-coefficient improves substantially if a rater is omitted, then this rater could benefit from more training on this item. The intended use of the MBT-G-AQS concerns decisions of whether subjects are below or above some specific level of adherence or quality. Consequently, the most relevant reliability estimate is absolute decisions (i.e., absolute G-coefficients; see Karterud et al., 2013 for a detailed discussion of this topic). Within the design of G-theory, several variance components can be disentangled in just one analysis (Shavelson & Webb, 1991). However, different designs depend on whether raters are crossed with patients (e.g., all raters rate all patients), whether they have unique raters nested within patients (independent groups of raters and patients), or whether the raters are considered to be random effects (to be representative of raters beyond themselves) or fixed effects, only to represent themselves (Shrout, 1998).

In the current research design, two therapy sessions from each of eight pairs of therapists were videotaped. Five raters rated all 16 sessions in this study. In the framework of G-theory (Shavelson & Webb, 1991), this implies a two-facet partially nested "(s:t) x r" design, where sessions (s) are nested within therapists (t), and raters (r) are crossed over sessions within therapists. The design is partially nested because the effect of a session (s) is both nested (within t) and crossed (over r). The two facets of observation give two differentiation variance components, the individual variance between therapists (t) and the systematic variance between sessions for each therapist (st). This makes three sources of instrumentation variance (error) that directly affects the reliability of the observed scores. These are 1) the rater effect (r) indicating the consistency of how much 'behavior' the raters see, averaged over therapists and sessions; 2) the interaction between raters and therapists (tr), indicating the raters' different rank ordering of the therapists; and 3) the unique rater–therapist–session interaction plus other unknown error variance (rst, + e). See Figure 2 for an illustration of the (s:t) x r design. In this design, sessions (s) cannot be separated from a therapist (t) and neither can the session–rater interaction (sr) be separated from the rater–session–therapist interaction. "By explicitly recognizing that multiple sources of random and true score variance exist and that

measures may have different reliabilities in different situations, GT has many advantages over classic true score theory" (Pedersen, 2008, p. 34).

3.1.6   Ethics

For the MBT-G sessions, patients received a description of the study and provided written informed consent, as did the therapists involved. The PDG recordings were part of the UPP project and were approved by the Regional Ethics Committee in Norway. The privacy ombudsman at Oslo University Hospital approved the MBT-G part of the study.

3.2 Paper II

This paper is a qualitative study of in-session therapy processes.

The material for this study was selected from the Quality Lab for Psychotherapy at Oslo University Hospital (http://www.psykoterapilab.no).

3.2.1   The Quality Lab for Psychotherapy

Karterud et al. (2013) reported a G coefficient (G-study) of .84 and .88 with seven raters for adherence and competence, respectively for MBT-I. This study formed the backbone to initiate the "Norwegian MBT quality lab", which later became the "Quality Lab for Psychotherapy". Its primary task is to assist local MBT programs with quality control of their treatment integrity. Local MBT programs deliver video recordings of therapy sessions to the lab. The two main tools for the lab are the MBT adherence and competence scale for individual MBT sessions (MBT-I-ACS) and the adherence and quality scale for group MBT sessions (MBT-G-AQS; Karterud, 2012, 2015; Karterud et al., 2013). The term competence was replaced with quality in the group scale. In addition to MBT-I-ACS ratings, the local MBT programs receive a clinical evaluation providing clinical guidance (e.g., regarding both alliance and technique). The individual MBT-I-ACS consists of 17 items (see appendix A). The item descriptions and rating procedures are presented in the "Manual for mentalization-based treatment" (Karterud & Bateman, 2010). The MBT lab has assessed MBT treatment integrity for research projects at the Psykiatrisk Klinik Roskilde in Denmark, the Department for Personality Psychiatry at Oslo University Hospital in Norway, and the Stockholm Centre for Dependency Disorders in Sweden.

3.2.2   Sample

The four most deviant (extreme) sessions were sampled from a total of 108 individual MBT sessions previously rated with the MBT-I-ACS. The data material in the qualitative analysis consisted of anonymous transcripts of these sessions. All four sessions were part of comprehensive MBT treatment programs. This means that 1) all patients suffered a PD in the borderline range and 2) the individual sessions were part of a broader program that also included psychoeducation (during the initial stage) and MBT-G therapy. At the time of the video recordings the treatment had lasted for various time periods ranging from 6–24 months.

Therapists

The four therapists were experienced psychotherapists with a mean age of 55. All four therapists had completed advanced training courses in MBT. All therapists received regular MBT (group) supervision.

Verbatim transcripts

In this study, the four sessions were transcribed. Personal data was altered and anonymized (i.e., names of friends and relatives, workplaces, toponyms).

MBT-I-ACS ratings

The ratings this study was based on were performed by the MBT lab and included both independent separate ratings and consensus study of video-recordings. In further analyses in the present study, authors EF and SK re-evaluated transcripts independently and by consensus. The rating procedure for the 17 different MBT-I-ACS is identical to that for the MBT-G-AQS.

Raters and reliability

The reliability between EF and SK was assessed based on 30 previous MBT-I-ACS ratings. The reliability was very high (mean value of absolute G coefficients) at .95 for adherence (range: .87 [Item 9]–1.0 [Item 15]) and .9 for quality (range: .82 [Items 2, 10, and 11]–.98 [Item 17].

3.2.3   Qualitative methods

After completing the ratings, the complete transcripts were searched for relevant excerpts to describe, exemplify, and illuminate the macro- and micro-processes where the therapist displayed a clear focus on the main goals of the treatment/session (e.g., to stimulate

mentalizing, to challenge unwarranted beliefs, to negotiate ruptures in the alliance, to handle aggression). Our epistemological stance regarding the present data is grounded in philosophical hermeneutics (Gadamer, 1975; Habermas, 1986; Schwandt, 2000) in which meaning is negotiated and understanding is interpretation that presupposes, ideally, the engagement of one's own biases and prejudices. The interpretation of a given text will change depending on the questions the interpreter asks of the text. Gadamer (1975) reconceptualized the hermeneutic circle as an iterative process through which a new understanding of a whole reality is developed by exploring the details of existence. The four sessions were analyzed using the framework provided by interpretative phenomenological analysis (IPA; Smith et al., 2009).

Interpretative phenomenological analysis

IPA is a recently developed qualitative approach that has rapidly become one of the best known and most commonly used qualitative methodologies in psychology. Between 1996 and 2008, 293 papers presenting empirical IPA studies were published (Smith, 2011). IPA is typically concerned with the detailed examination and interpretation of personal lived experience and one makes sense of that experience. IPA is an experiential psychological approach that draws inspiration from phenomenological philosophy and hermeneutic theory. In this spirit, IPA encourages researchers using the approach to engage with its theoretical and epistemological underpinnings. This is in line with Habermas (1986) and Ricoeur (1970, 1996), who provide the fundamental building blocks in our analytic approach (IPA).

Habermas is central in the hermeneutic tradition, which together with phenomenology comprises the backbone of IPA (Eatough & Smith, 2008, p. 194). Therefore, Habermas' (1986) three validity claims form the epistemological foundation for the current thesis. IPA is in the upper-left quadrant (I), and subjective sincerity or transparency is the key to applying this hermeneutic method for analyzing data. While hermeneutics constitute the theory and methodology of interpretation, the other central element of IPA is the study of structures of consciousness as experienced from the first-person point of view, phenomenology. When building theories from our observations and reflective processes, the following passage from Habermas underscores how we understood the phenomenological process in Paper II: Analytic insights, he argues, "possess validity for the analyst only after they have been accepted as knowledge by the analyst himself. For the empirical accuracy of general interpretations depends not on controlled observation and the subsequent communication

among investigators but rather on the accomplishment of self-reflection and subsequent communication between the investigator and his 'object'" (Habermas, 1986, p. 261). However, as we did not check our interpretations with the patients or the therapists, we did our best to (phenomenologically) inhabit their lived experience (Smith et al., 2009). Somewhat overlapping with mentalizing, one central structure of individual experience is intentionality or some theory of what motivates action. This presupposes researchers' ability to mentalize, and IPA is thus closely connected to the investigated material.

Such an analytic position becomes clearer when we introduce Ricoeur (1970, 1996), who suggested there are two kinds of hermeneutics: 1) Empathic interpretations are motivated by a desire to get as close to the meaning of a text as possible by trying to understand it "from within", just like mentalizing is also defined as understanding others from within and yourself from the outside. However, such interpretations focus on how (rather than why) something is experienced and presented. 2) Suspicious interpretations, as the term implies, have deep connections to a detective bureau but also classical psychoanalysis, where nothing is accepted at face value but is rather an expression of something hidden. If we were to place different research traditions on this continuum, then Grounded Theory, Q Methodology, Phenomenological Methods (e.g., IPA) lean towards empathic interpretation, while, Discourse Analysis, and Psychoanalytic Approaches lean towards suspicious interpretation. However, the idea of a continuum is misleading, as the entire point is the dialectic circle these forms of interpretation imply. However, these two ingredients can occur in various amounts in the larger cocktail (Smith, 2011).

Consequently, while MBT promotes a mentalizing stance, Ricoeur invites us into the hermeneutic circle through a Hegelian dialectic between understanding based on an "empathic stance" and explanations stemming from a "suspicious stance". Just like quantum theory, the concept of the "hermeneutic circle" acknowledges the impossibility of approaching a phenomenon without adopting a particular perspective or stance in relation to it. Without a standpoint, we would not be able to find meaning in what we examine. Therefore, we need to apply assumptions and ideas to begin to make sense of what we investigate. However, we avoid pretend mode and do not simply project our expectations onto a blank screen in the outside world and then find what we are looking for. In this interaction between our ideas and the world, our ideas evolve to accommodate what we have encountered. This implies an interdependency between the parts and the whole; the words

within this sentence are best interpreted on the basis of the sentence, the paragraph, and the larger dissertation it is part of, which again is part of a larger system, which again meets your mind and your ever-evolving evaluation of what on earth is going on when this sentence stops. Such progress is not always linear (or even curvilinear), and in meeting another subject the miracle of understanding is that no like-mindedness is necessary for recognition (Gadamer et al., 2004, pp. 309–310). IPA can be understood as a framework to underpin such a process or a way to structure and/or organize it.

IPA starts with probing the particular and ensures that generalizations are grounded in the idiographic details. However, in an attempt to create chain reactions from the particular and uniquely specific (e.g., like splitting the atom), IPA invites an intensive investigation of individuals' intrinsic psychology in the epistemological journey to universal laws and principles. Such a science is always aimed at a cautious climb up the ladder of generality, seeking universal structures but reaching them only via a painful, step-by-step approach (Eatough & Smith, 2008). Smith (2011) has developed a set of evaluative criteria for IPA that include a sustained focus on a particular aspect of experience, rich experiential data, assessment of the thematic structure using a measure of prevalence, careful elaboration of themes, and, of course, detailed, interpretative engagement with the material.

Hermeneutic circles thus encompass the deep insight also encountered in the double slit experiment (e.g., Penrose, 2006) in quantum theory—the fundamental interconnectedness of observer and observed. Thus, just as the knife cannot cut itself, we cannot free ourselves from our presumptions about the world.

> Rather, said Hegel, we must realize that thoughts are not merely a reflection on reality, but are also a movement of that very reality itself. Thought is a performance of that which it seeks to know, and not a simple mirror of something unrelated to itself. The mapmaker, the self, the thinking and knowing subject, is actually a product and a performance of that which it seeks to know and represent. (Wilber, 2000a, p. 59)

It follows that the scientific ideal in IPA is a transparent stance when interpreting information.

IPA is often applied to interviews in which the phenomenological interpretations are supported by a discourse with the subject and typically recommends researching few in-depth interviews. Paper II includes four transcripts, but the sessions do not count as interviews.

However, despite this limitation, we believe that clinical expertise allows for close investigation of such transcripts, especially in terms of the theories, experience, and treatment method. IPA seemed ideal in terms of a fundamental investigation of phenomena such as alliance, epistemic trust, and strategic competence. The IPA analysis in Paper II involved five steps (see the published paper for details):

1) The four sessions were transcribed and studied in detail.
2) We phenomenologically investigated the therapeutic alliance (goals, tasks, and personal bond).
3) Four emergent themes began to appear.
4) These themes were debated in light of psychotherapy research.
5) The authors decided on the major themes in the investigated sessions.

3.2.4 Reflexivity

It is important to be transparent about our interpretations (Finlay & Gough, 2008). In Paper II, the first (EF), second (SK), and last author (ES) were most involved with the transcripts and the IPA. The third (MK) and fourth (EK) authors were part of the later analysis and the overall interpretations in the manuscript. EF originally had the idea of merging the CF and EST approaches within this field when he became a rater at the Quality Lab for Psychotherapy in (January) 2013. As a student, he had rated hundreds of therapy sessions from a previous RCT (Svartberg et al., 2005) and trained other raters in the project headed by Leigh McCullough at Modum Bad (where he discovered that he preferred the alliance in psychodynamic therapy above that in cognitive therapy). EF was strongly influenced by McCullough, but his psychodynamic orientation is perhaps closest to object relations theory (e.g., Bion, Kohut, Fairbairn, Winnicott, and Guntrip) and Eastern philosophy traditions (especially Buddhism). This would have influenced his preference for a highly open-ended approach (e.g., listening before talking and being non-judgmental and accepting) colored by a well-timed transference of knowledge. Therefore, some time elapsed before EF accepted the battling stance evidently associated with highly rated MBT. However, EF also feels very much at home within the mentalization tradition and finds that an active therapeutic stance seems natural when treating PDs. EF has also been a professional musician since a young age, and the pedagogic stance embodied by the eminent pianist Sir András Schiff comes close to his ideal of such an approach—deeply listening to each student to try to bring forth their unique quality. EF has also been deeply impacted by being a student of the Diamond

Approach from 2004–2018. This approach advocates that one size fits none and that each meeting between two individuals is unique. When analyzing the four selected sessions in Paper II, it became evident that EF has a tendency to prefer transparent therapists who lean towards self-disclosure, such as the therapist in Session B who uses his bond with Elsa to nudge her into group therapy (Paper II). ES shares many of the same preferences and comes from a psychoanalytical background, being most interested in the fundamental mechanisms in psychotherapy. It was also very beneficial for the analysis that one author (ES) is not in the field of MBT. The rest of the authors are all experts in MBT. Paper II was ambitious in terms of trying to convey something deep, while being firmly rooted in empirical observations.

The author of this dissertation may have been influenced by being co-author on the new Nordic MBT manual (Karterud et al., 2020). One example of this is that the proposed mechanism of change in MBT

> is to stabilize mentalizing in certain focus areas to create a psychic buffer between affect and behavior to foster affect regulation, reduce impulsivity, and promote functional supportive relationships. This is reached by employing "contrary moves" to create more flexibility in using the different poles of mentalizing. (Taubner & Volkert, 2019, p. 8; Volkert et al., 2019, p. 25)

Such "contrary moves" are not considered specific MBT interventions in the new Nordic MBT manual (Karterud et al., 2020)—not because they fail to be clinically relevant, but because they seem hard to define (e.g., evaluate for adherence and quality). Indeed, the reader may be able to easily imagine numerous psychotherapeutic interventions (strategies) that are not contrary moves, and many of them likely also characterize manualized MBT. Further, I have thought that the field of MBT may be ill served by proposing such abstract theories, as if they are the sole solution, and also somewhat making mentalizing the cornerstone of all psychotherapies, typically arguing that neuroscience "strongly suggest that mentalizing is an evolutionarily prewired capacity" (Luyten et al., 2020a, p. 298). I fully agree that "mentalizing may be commonly found as a factor associated with recovery in a range of psychotherapies" (Luyten et al., 2020a, p. 299), but perhaps due to the absence of a clear definition of such conceptions, I fail to understand why it needs to be stated that mentalizing may be a change agent also when the therapeutic discourse is not focused on mentalizing:

> [I]nterventions that may not explicitly focus on improving mentalizing nevertheless may be effective in reducing mental health problems as they may foster mentalizing and salutogenesis in particular through different routes. (Luyten et al., 2020a, p. 299)

Further, in terms of working with the concept of the alliance, one needs to be as aware of one's preconceptions as possible. The alliance, whether interpreted by observer, therapist, or patient, needs to be seen in the context of the features of purposive, collaborative work specific to the given treatment (Horvath, 2018). Consequently, the specific treatment will inform us how we are supposed to judge the alliance, a consensus we may or may not adhere to, depending on our preferences. An excellent alliance in CBT may not be very similar to an equally strong working relationship in psychoanalysis. Hence, the alliance is a term that needs to be defined clearly in order to inform the reader exactly what kind of behavior it denotes. This is also true in terms of time, for instance, an alliance rupture could denote events between sessions, within a whole session, between segments of the session, between grammatical units, or within utterances (Horvath, 2018). As all these time frames represents a legitimate and meaningful conceptualization of disruption in the alliance, we tried to bring awareness to such perspectives in the IPA.

3.2.5 Ethics

All patients gave written informed consent to participate in the study. The procedures for recording and transcribing the sessions were approved by the privacy ombudsman at Oslo University Hospital. One session was approved by the Psychiatric Research Unit at the Psychiatric Clinic in Roskilde, Denmark. The other was from an earlier RCT study (Mentalization-Based Treatment for Dual Diagnosis) conducted at the Center for Dependency Disorders at the Stockholm County Council and Karolinska Institutet in Sweden. This project was approved by the Regional Ethical Review Board in Stockholm. The therapists consented to this publication. Interpreting the style of individual therapists, perhaps especially in the low rated cases, raises several ethical concerns. It may weaken the therapist's self-esteem and thus impact future treatments. It may also make the therapist reject the MBT model and perform therapy in a manner where they are not evaluated by strict raters of adherence and competence/quality. We strived to communicate that highly rated and poorly rated MBT does not mean good and bad therapy but that such ratings only concern how much MBT the relevant sessions contain. This is a challenging task, which is also a continual challenge for the Quality Lab for Psychotherapy. Achieving a very high rating may cause similar problems, and the same message was communicated in these instances (i.e., high ratings only imply that the therapy resembles the ideal in MBT).

3.3 Paper III

This paper is a quantitative study of longitudinal data.

3.3.1 Subjects

The studied sample comprised 155 patients treated on a regular basis with MBT on a specialist mental health service level during 2009–2016.

Mentalization-based treatment

The MBT program was an intensive, long-term outpatient treatment in accordance with MBT manuals (Karterud, 2011, 2012; Karterud & Bateman, 2010). The first year included weekly individual and group therapy sessions and psychoeducational group session (12 sessions). The frequency of individual therapy was gradually reduced in the second and third years, while group sessions continued throughout treatment. Treatment had an upper time limitation of 36 months.

Therapists

The team was multidisciplinary (three psychiatric nurses, three psychiatrists, an art therapist, a physiotherapist, a social worker, and two psychologists). Eight were qualified group analysts, one was qualified in psychoanalysis, and one was qualified in individual psychodynamic psychotherapy; 67% are female, and the mean age (year 2009) was 53 (SD = 9). Other individual therapists involved in the research period were different resident doctors and psychologists in training. All had basic MBT training and attended regular weekly video-based supervision by qualified MBT supervisors.

3.3.2 Therapist mentalization-based treatment fidelity

The present study includes measures for treatment fidelity (ratings of MBT adherence and competence) in all three papers. MBT adherence and competence was assessed using videorecorded therapy sessions, the MBT Adherence and Competence Scale (Karterud et al., 2013), and the Adherence and Competence Scale for Mentalization-based Group Therapy (Karterud, 2015). On a 1–7 scale, a score of four or higher indicates adequate MBT adherence/competence. In 2013–2015, five raters evaluated 19 individual sessions (eight therapists) and nine group sessions in the program. For individual therapists, the mean adherence level was 4.7 (SD = 1.2) and the mean MBT competence level was 4.4 (SD = 1.2) (Kvarstein et al., 2015). For group therapists, the mean adherence level was 5.1 (SD = 1.37)

and the mean competence level was 4.9 (SD = 1.3) (Kvarstein et al., 2020). This is comparable to a recent RCT study of MBT in groups for adolescents with BPD (Beck et al., 2020). In that study with experienced and motivated therapists, the mean overall adherence score was 5.47 (SD = 0.80) and the mean overall competence was 5.53 (SD = 1.10).

3.3.3   Baseline assessment

Assessment of diagnoses and personality functioning at the start of treatment (baseline)
Diagnoses were based on the Mini International Neuropsychiatric Interview (MINI) version 4.4 for DSM Axis-I diagnosis (Sheehan et al., 1998) and the Structured Clinical Interview for DSM Disorders (SCID-II) for DSM Axis-II diagnosis (First et al., 1994) (DSM-IV). Experienced (10–20 years of practice) and specifically trained clinical staff performed the MINI and the SCID-II interviews.

Aspects of personality functioning were measured using the Severity Indices of Personality Problems (SIPP-118), a 118 item self-report questionnaire aimed at measuring five core domains of personality pathology—self control, identity integration, responsibility, relational functioning, and social concordance (Verheul et al., 2008). High scores indicate better adaptive functioning, whereas lower scores indicate more maladaptive personality functioning. The SIPP subscales have generally yielded adequate to strong internal consistency in PD samples, with α scores ranging from .62–.89 (Feenstra et al., 2011; Verheul et al., 2008). A Norwegian replication of the original Dutch study found good cross-national validity of the SIPP-118 (Arnevik et al., 2009a). Further, all facets of SIPP have good discriminative properties with respect to differentiating between a nonclinical sample, a clinical sample without PD, and a clinical PD sample (Pedersen et al., 2017a).

Quality of Life (QoL) was measured at the start and end of treatment using the EuroQol (EQ-5D). The EQ-5D is a self-report questionnaire that provides a simple method to measure health problems in five dimensions—mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. EQ-5D is a useful tool to assess QoL in patients with BPD (Van Asselt et al., 2009) and is sensitive to change in patients with PDs. The QoL index score as measured with the EQ-5D is expressed as a single index score ranging from 0.33 (worst imaginable health state) to 1.00 (best imaginable health state). According to reports on the QoL index score with BPD patients, the score ranges from .44–.57 (Laurenssen et al., 2016). One Dutch

study included 403 BPD patients and found a mean QoL index score of .48, which is comparable to that of patients with severe physical conditions, such as stroke or Parkinson's disease (Laurenssen et al., 2016). The mean QoL index score of the general population in Western societies ranges from .83–.87 (Saarni et al., 2007).

Interpersonal problems were assessed using the Circumplex of Interpersonal Problems (CIP; Pedersen, 2002). CIP is a 48-item version of the Inventory of Interpersonal Problems-Circumplex version (IIP-C) self-report questionnaire (Alden et al., 1990). Severity is rated on a 0–4 scale (0: not at all; 4: extremely). The mean sum score (CIP) correlates r = .99 with the original IIP-C sum score (Pedersen, 2002). The reliability of CIP is high ((four-day test–retest coefficient [ICC, 2.1], r = .96, 95% CI; .93–.98; Pedersen et al., 2011). In a non-clinical Norwegian sample, mean CIP scores were 0.5 (SD = 0.3) (Pedersen, 2002). Including one standard deviation, the clinical/non-clinical CIP cut-off score is 0.8. CIP scores of 1.7 and above indicate severe interpersonal distress, scores of 1.3–1.6 indicate significant to moderate interpersonal distress, and scores below 1.2 indicate insignificant to mild interpersonal distress (Pedersen, 2002).

3.3.4 Measure for therapeutic alliance: Working Alliance Inventory

The vast majority of empirical studies of the alliance reference Bordin's writings as a way of defining the concept (Horvath, 2018). There are several ways to assess the concept of working alliance (Bordin, 1979, 1994), such as by rating the alliance using the Working Alliance Inventory–Observer version (Darchuk et al., 2000). In Paper III, the Working Alliance Inventory–Short Revised (WAI-SR; Hatcher & Gillaspy, 2006) was used to measure the patient-rated alliance. Although the alliance is not reducible to the patient's experience of it, the patient's experience is important in understanding the relationship between alliance and outcome. In fact, patient-rated alliance has been a consistent predictor of outcome for decades (Flückiger et al., 2018; Horvath & Symonds, 1991). Proposed by Bordin (1979), the original WAI is a 36-item measure that assesses three aspects of the therapeutic relationship: (a) the bond between patient and therapist, (b) the extent to which the patient and therapist agree on the goals of treatment, and (c) the extent to which the patient and therapist see the tasks of therapy as relevant. All 36 items may be aggregated to create a total score, with high scores reflecting strong alliances. The WAI has been shown to have good internal consistency (α = .93) and adequate convergent and predictive validity (Horvath & Greenberg, 1989). There is an abundance of scales to measure the working alliance (Falkenström et al., 2015), but the

WAI has the clearest conceptual foundation, as it was developed based on Bordin's (1979) conceptualization of the alliance as composed of agreement on goals and tasks and supported by bonds. The WAI includes no items referring to specific treatment methods. In short, these measures give a general overall reading of the state of the working alliance at the session level.

> Although there is some evidence that these measures tap discernible dimensions of alliance (e.g., Hatcher & Gillaspy, 2006), a compelling argument can be made that these measures are like room thermometers in that they give an overall reading of the quality of the working alliance without being very localized or specific about it."
> (Hatcher, 2010, p. 15)

Importantly, Falkenström et al. (2015) reported that the intercorrelation between the task and goal factors in their three-factor model was high and concluded that a two-factor structure where the task and goal factors are collapsed into one is psychometrically more defensible than a three-factor structure. This makes two main factors of interest to investigate 1) bonds and 2) tasks and goals. The WAI-SR is a short form of the patient version of the (WAI; Horvath & Greenberg, 1989) and consists of 12 items rated on a seven-point scale ranging from "never" to "always." Based on Bordin's (1979) conceptualization of the working alliance, it has three subscales, goals, tasks, and bonds, with four items for each. The short version consists of the items with the highest load on each of the three subscales. Each item is rated on a seven-point scale, with higher scores indicating better alliance. The WAI-SR has been found to have good psychometric properties (Munder et al., 2010).

In our study, we investigated the three subscales in separate, independent models. Scores above four indicate satisfactory alliance. The WAI-SR scores are presented in Table 2. In the first period of our study, patients received the "old WAI version" (n = 34 patients). From June 2012, all patients received the "new WAI version" (n = 71). Eighty patients in the current investigation have longitudinal data including for both versions of WAI. There are some minor changes between the old and new versions of WAI. However, most variance between the two versions is within and not between the three alliance categories; at a conceptual level, the two WAI versions measure the same in terms of tasks, goals, and bonds.

3.3.5 Main measure for clinical outcome: Global assessment of functioning

The first standardized and broadly used instrument for assessing patients' overall

mental health was introduced by Luborsky in 1962 when he developed the Health-Sickness Rating Scale (HSRS; Luborsky & Bachrach, 1974). Some decade later, Endicott et al. (1976) modified the original instrument, which resulted in the Global Assessment Scale (GAS). Both the HSRS and the GAS are single 100-point rating scales reflecting overall functioning from 1 to 100, where 100 would be the hypothetically sickest patient imaginable and 1 the hypothetically healthiest individual. GAF is a widely used measure; it is observer-rated and is a composite variable reflecting both symptoms and functioning, which can be seen as an advantage. GAF includes both symptoms and functioning and reports the lowest value of the two. In Paper III, GAF was chosen as parameter for clinical outcomes, as the observer-rated GAF provides a composite score combining social and symptom-related dysfunction (0–100 scale, Axis V, DSM-IV) (Pedersen et al., 2018). Higher GAF scores indicate better overall psychosocial functioning, and a score of 60 represents the cut-off level between mild/no impairment and moderate/severe impairment. Conventional interpretations of severity indicated by GAF scores are as follows: mild (61–70), moderate (51–60), and severe (41–50) (Pedersen et al., 2018). Staff therapists were trained (GAF assessment courses within the Norwegian Network for Personality Disorders) and then performed GAF evaluations. The reliability of the applied method for GAF assessments was tested in 1998 (staff consensus scores) and 2001 (independent scores) (Pedersen et al., 2007). Clinical vignettes were scored by staff consensus in eight different treatment units (including the studied treatment unit) by 58 staff members. Reliability for consensus scores was high (ICC 2.1, single measure, absolute agreement definition: 0.94, 95% CI 0.85–0.98). Adequate reliability and validity of the GAF was reported in a publication by Pedersen et al. (2018). Consistency of GAF scores across units and raters was also high (generalizability coefficients of absolute decision (the score) range .86–.95) (Pedersen et al., 2007). On treatment termination, 59% had scores of 60 or higher (mean GAF end score of 70, SD = 7), and 41% had end scores less than 60 (mean GAF score of 51, SD = 7) in our sample (N = 155). In this study, the sample was divided into two subgroups:

1) Patients with clinical outcomes within a clinical range (GAF below 60) at the end of treatment
2) Patients with clinical outcomes within a non-clinical range (GAF equal or above 60) at the end of treatment

3.3.6 Supplementary clinical outcome measures

As our main indicator of clinical outcome was observer-rated, we included two supplementary patient-rated measures to assess functioning and symptom distress, The Brief Symptom Inventory 18 (BSI-18) and the Work and Social Adjustment Scale (WSAS). In Paper III, three variables were chosen to reflect clinical outcome—two self-report measures (BSI-18 and WSAS) and the observer-based assessment of GAF. GAF improvement trends in the sample corresponded with improvement trends for the two self-reports. We therefore chose the variable GAF as a main measure to indicate improvement because it had the advantage of combining both symptoms and functioning, and we had more complete GAF data for the final assessment.

Work and Social Adjustment Scale

The WSAS is an outcome measure assessing the degree of functional impairment (i.e., work, social, and private leisure activities and home, work, and social relations) (Pedersen et al., 2017b). Each of the items is rated on a nine-point Likert scale from "not at all" (0) to "severely impaired" (8). The total sum score of all items ranges from 0–40, where higher scores indicate more distress. Scores below 10 indicate a subclinical population, scores from 10–20 indicate significant but not severe functional impairment, and scores above 20 indicate moderately severe to severe impairment (Mundt et al., 2002). Patients also reported current status regarding work functioning, in terms of how many months they participated in more than 50% work or study during the previous year.

The Brief Symptom Inventory 18

The BSI-18 is a self-report questionnaire assessing symptom distress (depression, somatization, and anxiety on a 0–4 format scale; 0: "not at all", 4: "extremely"). BSI-18 includes an overall severity index, the mean sum score (BSI). The BSI-18 is derived from the 53-item BSI, a shortened form of the (SCL-90-R; Derogatis, 2000). The BSI-18 applies the same clinical case rule originally developed for the SCL-90-R. A conservative cut-off for clinical/non-clinical ranges of severity (sum score 0.8) is based on Norwegian sample norms and patient samples (Pedersen & Karterud, 2004). The BSI-18 was administered to all patients in MBT.

3.3.7 Data analysis

Mixed model analysis

Longitudinal data characteristically imply repeated observations of the same individual. Such repeated observations cannot be regarded as independent observations (Fitzmaurice et al., 2008). The sample in Paper III also had unbalanced data with different numbers of assessments per patient. Characteristically, mixed models do not require that all subjects have equal numbers of assessments or that the time intervals between assessments are constant (Norusis, 2008). Unlike somewhat simpler methods, such as repeated measures ANOVA, Linear Mixed Models (LMM; Singer & Willett, 2003) allows for the inclusion of cases with missing values and not only patients with complete datasets. The LMM method was chosen to maximize the use of available patient data. The modelling procedure starts with a model where the time interactions and random effects are not specified. This first unspecified model provides an estimate (fixed effect), corresponding residual variation, and log likelihood estimations of the goodness of model fit (Singer & Willett, 2003). Further specification of a linear model is a stepwise procedure adding a time interaction (fixed effect) and then random effects in accordance with the principle of achieving the best possible goodness of model fit (see Paper III for equations and details).

The three WAI-S/SR subscales were the dependent variables. The sample was modelled with the GAF ratings in two subgroups according to outcomes (scores at discharge) above or lower than clinical/nonclinical cut-off levels (60) as predictor. The LMM included longitudinal change of WAI-S/SR subscales and variation associated with the dichotomous GAF variable as a predictor, as well as a moderator interaction combining the dichotomous GAF variable and patient factors.

All included patients had at least one assessment (even patients with only one assessment can be included in the analyses) (Singer & Willett, 2003), and the mean number of WAI-S/SR assessments was 3.2 (SD = 1.8, range 1–9). The sample had unbalanced data with different numbers of assessments per patient. As LMM incorporates unbalanced data and uses all available data for each individual trajectory, we did not use imputation methods for missing data. A variable counting numbers of assessment points for each individual captured a relevant missing data pattern. To investigate the effect of this missing data pattern on the outcomes, the variable was added as a predictor in all three working alliance subscale models (Hedeker & Gibbons, 1997). Analyses showed poorer initial alliance ratings for patients with fewer assessments ($p < 0.05$ for all working alliance subscales) but no significant effect of the variable on alliance development over time ($p > 0.05$ for all WAI subscales).

3.3.8 Ethics

All research was performed on anonymous clinical data from an anonymous research database with approved procedures. All patients gave written informed consent to participate in the research. The treatment unit collected clinical data, which was registered in an anonymous database administrated by Oslo University Hospital. Procedures for data collection ensured that participating individuals could not be identified. Data security systems were approved by the Data Protection Official at Oslo University Hospital. Because the data was anonymous, ethical approval was not required from the Regional Committee for Medical Research and Ethics.

4 Results

4.1 Paper I

The results show high reliability for both adherence and quality (competence). The mean absolute G-coefficient for adherence was .86 (range .63–.97) and was .88 for quality (range .64–.96). The reliability for overall adherence (.97) and quality (.96) ratings are both high. The nine group-specific items (items 1–9) displayed very high reliability for both adherence (range .83–.95) and quality (range .78–.96). Further, the residual variance for the overall quality score was very low (17%), and there was complete agreement among the raters on frequency/adherence and the ranking of therapists. The results show only minor differences between relative and absolute G-coefficients (raters agree as much on exact scores as on the ranking of the sessions). Table 3 shows the grand mean and standard deviation of scores across all raters and sessions and G-coefficients for all ratings of all items. Reliability was very high to excellent for the entire scale and for single items (items with low absolute G-coefficients had very low variance, and the reliability is therefore also acceptable/high for these items), indicating reliable assessment of the specific aspects of competence and adherence for MBT-G. The nine group-specific items had as high reliability as the nongroup-specific interventions. Items 9, 14, 16, and 18 have a low frequency (adherence) rating, while Items 11 and 14 are rated often.

For some of the items the reliability would increase notably if one of the raters (different items for different raters) was omitted in the G-study. Importantly, the overall ratings were the most robust items for a decreasing number of raters. Deleting the least reliable rater from the overall ratings would only slightly increase the reliability for these two items (+0.01). Further, the absolute G-coefficient for Item 8 (adherence) would increase from .83 to .89 if rater number 5 was excluded. If rater 2 was omitted, the absolute G-coefficient of adherence would increase for Item 9 (from .84 to .89) and for Item 11 (from .80 to .87). Omitting rater 4 from the study would increase the absolute G-coefficient of adherence for Item 18 (from .88 to .93). Excluding raters 2 and 4 from the study would also increase the absolute G-coefficient of quality for Items 18 (from .84 to .87) and 19 (from .88 to .93). This means that different raters had areas in which their understanding of the scale deviated from the "norm" but signals that variance is not systematic in terms of one rater being consistently worse than the others. Regarding adherence (frequency), the present study displays very high variance (sum of variances) for Item 11 (21.77), Item 17 (28.76), and Item 19 (27.18) and very low variance

for Item 14 (1.4), Item 16 (.77), and Item 18 (.72). The reliability for Item 16 (and Items 14 and 18) is very good considered such low variance. On Item 11, the raters differ to a greater degree, even though the G-coefficient here is high (.80). Regarding quality, our G-study indicates a low sum of variances for Item 1 (.65), Item 10 (.99), Item 12 (.92), and Item 15 (.67), with the highest variance for Item 2 (3.13), Item 3 (2.98), and Item 9 (2.85). According to the quality ratings of Item 8, the therapists vary greatly from session to session (the S:T-variance = 1.69; sum of variances = 1.69). Additionally, for Items 8 and 9 there are no individual differences (variance component from therapists = .0).

One finding not emphasized in Paper I was that on a descriptive level, there was a noteworthy difference between treatments (PDG versus MBT-G) in quality/competence but not adherence. This signals a significant correlation between adherence and quality/competence but also that the MBT-G-AQS had some discriminant validity. The structuring elements of the MBT-G-AQS (Items 1, 2, and 3) were the major difference from PDG. Table 4 shows that Items 2, 3, and 11 display the largest difference between PDG and MBT-G. Item 9 had a higher prevalence in PDG (M = 4) than in MBT-G (M = 0).

4.2 Paper II

Embedded alliance represents a marriage of the common and the specific psychotherapeutic factors, and its importance is one of the major conclusions of this thesis. Paper II was conceived from the viewpoint of the embedded alliance, but primarily described findings in common factor terminology. As mentioned, this was largely in order to avoid a circular logic. Further, one of the core research questions was whether the differences between highly rated and low rated MBT could be explained (in the language of) CFs. However, as this thesis concludes that the alliance is largely a product of the applied method (MBT), it seems reasonable to present the results from the viewpoint of the embedded alliance. I will first present the findings from Paper II, and then I will use the 17 MBT interventions (MBT-I-ACS) as a conceptual backbone for describing how the specific technique (MBT) fosters alliance (embedded alliance).

4.2.1   The four themes constituting battles of the comfort zone

The four themes that emerged from the IPA were labelled: 1) alliance, 2) strategic competence, 3) quality, and 4) battles of the comfort zone. Our observations indicated these were related in the following manner. All four identified themes seemed to interact and

reinforce each other. For instance, in the highly rated MBT therapies, therapeutic alliance was fostered by battles of the comfort zone, quality, and strategy. As long as there seemed to be an adequate alliance, therapists nurtured the alliance through battles of the comfort zone. When therapists competently challenged problematic patterns, despite disclosing discomfort, they seemed to fortify the alliance. This appeared to create somewhat of a snowball effect, such that the stronger the alliance the more the therapist could challenge the patient, thus being able to foster further enforcement of the alliance. The epistemic trust seemed to grow stronger through the experience of being seen and by having obtained (significant) new understanding about oneself as a product of the therapist having dared to follow the patient through painful terrain. Quality partially overlapped with strategic competence but could also be seen as a measure of the alliance strength. In highly rated MBT, the three first identified themes (alliance, strategic competence, and quality) worked together in such a way that the patients were brought out of their comfort zone, resulting in beneficial therapeutic work.

Let us first summarize the four sessions, before we explore different intervention types and how they informed our interpretation (IPA). In the summaries of the four sessions, in order to highlight the interconnectedness of MBT and the CFs (embedded alliance), the MBT-I-ACS ratings (not published in Paper II) will also be presented.

4.2.2 Summary of session A (Elsa)

This session concerns a female patient in her fifties who also receives treatment for heroin addiction. She starts talking about a funeral she helped arrange. The therapist challenges her self-devaluation and the patient eventually comes to acknowledge the good work she did for others and herself with the funeral. In the next sequence, the therapist comments "… and then you have not been in the group therapy, have you?" Her absence from the group is the main theme for the rest of the session, alternating with worries about her son's suicidal thoughts and the relationship to some friends. The patient strongly dislikes the group being a theme in the session: "No, that damned group. I hoped that you would forget about it". Initially, she resists, but slowly they approach the underlying painful theme of returning to someone who has hurt her. She promises to go to the next group session. In this session, "battles of the comfort zone" emerged in the therapist's effort to sustain a focus on mental states, as did the use of the positive bond to challenge the patient's massive resistance to the theme.

This session received a rating of 7 (excellent) for both adherence and quality. The number of interventions that adhered to the manual was high (N = 160), and the percentage of MBT interventions relative to all interventions was very high (85%).

4.2.3   Summary of session B (Maria)

This session concerns a female patient in her early thirties who starts with her hesitations about starting her college education at the same time as being in treatment. Lately, she has felt much better. However, the chronic feeling of emptiness is still present. The therapist suggests exploring this theme: "When did you experience that last time?" After the last group session, she says, and the group theme fills most of the remainder of the session. She is fairly new to the group and admits being irritated with senior group mates who have been reluctant to include her. The therapist explores in detail exactly what it was that provoked her. Gradually, a picture of a patient who has harbored strong resistance to the group therapy component emerges. When she eventually turned up, she was met with skepticism. She felt an urge to leave the group, thinking "fuck you". The therapist asks if some of her thoughts and feelings could be shared with the group. But she feels strangers should have no access to her inner life! This theme resonates with other relationships in her life. She has become rather lonesome. Her emptiness after group sessions is explored even more. The therapist asks, "Talking about it here, how does that affect you?" She replies, "Irritated/annoyed, really irritated/annoyed". She vents her feelings to the therapist, whom she feels is pushing her. It becomes a relief to have said this, and her feeling of being different and lonesome fills the last part of the session, now with tears and sadness. This session received a rating of 7 (excellent) for adherence and quality. There was a high number of interventions that adhered to the manual N = 193), and the percentage of MBT interventions relative to all interventions was high (76%).

4.2.4   Summary of session C (Diane)

This session concerns a female patient in her late twenties who attacks the therapist from the very beginning: "I was angry with you the last session. You said I was irritated. I was not; I was angry." The therapist seems to be taken by surprise and quickly becomes defensive. He has difficulty understanding the patient's point of view while also excusing his own behavior. The patient takes the initiative and talks on about several things, often in a pseudo-mentalizing way. The talk is about how the patient feels under pressure at work, about a good friend who does not understand her, about the patient's relationship with her mother (a theme the therapist is bringing in), about her general sense of not being understood (from within)

and of being judged from the outside, and about the difficult task of writing a study paper for the next day. The patient treats the therapist in a top-down manner, and the therapist succumbs. The therapist loses his mentalizing stance and moves into a kind of passive listening and supportive therapy. The therapist is harshly treated by the patient in the first quarter of the session and he submits to the patient's dominant style. Towards the end of the session, the therapist suggests an extra session to deal with the patient's manifold problems; this seems like reaction formation. The patient turns down the offer, saying that her problem actually is her limited time to write a study paper and that it will not help to talk to the therapist. In this session, the therapist did not battle the comfort zone and abandoned the main therapeutic project and goal (Theme 1). This session did not reveal relevant therapeutic work. It is quite possibly the therapist's own wish for a good, pleasant transference challenged the application of a focused technique and overall strategy. The session received a rating of 2 (poor) for adherence and quality. The number of interventions that adhered to the manual were few (N = 50), and the percentage of MBT interventions relative to all interventions was low (38%).

4.2.5 Summary of session D (Monica)

This session concerns a female patient in her early twenties who has resumed therapy after having missed a number of sessions due to a traumatic sexual assault five weeks earlier. The session is for the large part educative and counselling and is focused on practical issues regarding the patient's current life situation and how she deals with it, ways of taking care of herself, legal actions in relation to the assault, economic issues, advice on medication, etc. The therapist behaves in a caring and warm way and seems sincerely interested in the patient's situation. The therapist talks a great deal, leaving little room for the patient to speak about what she thinks and feels in the moment. The focus on mental states is largely left out of the session. In all, the session reveals a supportive kind of therapy with very few MBT interventions. Towards the end, the therapist takes up the subject of what kind of dress the patient is going to buy for the season and steers the conversation to superficial topics. The session ends 5 minutes before regular time, as the therapist and patient have no more to say. This session received a rating of 2 (poor) for adherence and quality. The number of interventions that adhered to the manual were few (N = 52), and the percentage of MBT interventions relative to all interventions was low (19%).

4.2.6 Quantitative characteristics of the MBT-I-ACS ratings

The mean number of ratings of highly rated MBT was 177 (80% of the total interventions), while low-rated MBT sessions averaged 51 ratings (28% of total interventions). Table 7 and Figure 7 show a large difference in numbers of MBT ratings for highly rated and low-rated MBT. Note that each intervention might receive several MBT ratings. Tables 5, 6, and 7 present different quantitative aspects important for our basic understanding of what is going on in these sessions. Figure 8 displays the mean adherence profile for high- versus low-rated MBT. The most frequent intervention for highly rated MBT was Item 17, "Integrating experiences from concurrent group therapy". In the low-rated MBT sessions, Item 3, "Challenging unwarranted beliefs", was absent, while in the highly rated sessions, Item 3 averaged 6.5. In summary, the highly rated MBT therapists employed far more interventions, especially Items 3 and 17. The low-rated MBT therapists mainly used Item 16, "Monitoring own understanding and correcting misunderstandings".

4.2.7   Results presented through the lens of embedded alliance

As we remember from the introduction, the term embedded alliance seems to adequately address the working alliance (e.g., agreement on tasks and goals) in line with the therapeutic model (theory of pathology and change). We have also seen that the 17 ingredients in individual MBT can be viewed as strategies to achieve a good alliance despite dysfunctional dyadic patterns. Paper II found that MBT can be described without using the specific concepts in MBT, which also avoided a circular logic. However, as one overall aim of the current thesis (the fourth question) is to put the puzzle of measuring MBT together, in the following I will show how the results from Paper II can be understood in terms of embedded alliance. For this purpose, the 17 items of individual MBT will provide structure, as I present those items that seem particularly important for the overall thesis.

4.2.8   Items 1, 4, and 6: Tailoring the treatment to the patient

MBT sessions of high quality display a good working alliance (the bond part, warmth, being part of Item 1) and high therapist competence and flexibility in tailoring the therapy for the specific patient/situation. The ability to adopt and maintain a mentalizing stance at the maximum of the patient's current mental and emotional capacity would be the chamber pitch of well-tempered MBT (Items 4, 5, and 6). MBT requires that the therapist recognize the patients' difficulties in mentalizing and try to join the process to amend this. Whenever the patient outlines their inner process, this is an opportunity for the therapist to focus on, train, and teach mentalizing. One patient (Diane) tries to mentalize in her first and second utterance

in Table 9. However, the therapist does not follow up on this invitation, presumably because he has lost his own mentalizing ability due to the patient saying she was angry with him in their last session. In session A, it only takes about 2 minutes from the session start before the patient is emotionally engaged in trying to understand herself. This hallmark high alliance is displayed by cultivating the core principle of treatment (mentalizing). This attests to the therapist's ability to combine Item 1 ("Engagement, interest, and warmth"), Item 2 ("Exploration, curiosity, and a not-knowing stance"), Item 4 ("Adaptation to mentalizing capacity"), and Item 6 ("Stimulating mentalization through the process"), as defined by the manual for MBT-I (Karterud & Bateman, 2010). This ability to initiate and join the patient in her mentalizing process also unlocks her epistemological trust. Early in the session (5 minutes), she states: "Oh, damn, what a skillful therapist you are! Thanks." The patient may not agree with the therapist or be able to understand all his views, but she is open to them, respects his opinion, and his statements make her investigate her own mental processes with sincere interest. The manual simply says: "The most important sign of a successful MBT session is that the patient gets involved in a mentalizing discourse" (Karterud & Bateman, 2010, p. 44). Paper II found that MBT with a high rating indicated active therapists who succeeded in engaging the patient in a mentalizing discourse, a sign of embedded alliance in MBT. Let us take a closer look at what this mentalizing process may look like.

4.2.9   Item 2: Mentalizing stance

As session A and B come to an end, it is with a distinct sense that the patients are left with alternative perspectives and improved mentalizing capacity (sign of high alliance). This change comes about because the therapists remain steadily anchored in an inquisitive but calm and open state of mind. All details the patient brings forth are seen as pieces of a larger puzzle, and the therapist advocates detailed accounts of experiences rather than explanations, in line with guidelines for MBT (e.g., Bateman & Fonagy, 2006). Still, there is a paradox in a "not-knowing stance" because the therapist applies all their knowledge and theories about the patient to guide the process (being steadfast to the project) through a practice of open-ended scrutiny. However, it is an art and a question of the therapist's best understanding to choose where to ask for and explore different alternatives. Early in Session B, the patient says: "Yes. Yes, yes, and I am also prepared that, if it should be, that I cannot, so if it should be, that, that my teacher does not want to give me dispensation, then I am fully aware that I will have to drop the education." The therapist could have but does not take this opportunity to explore if there could be alternatives to "drop the education", which seems somewhat nonattending to

the outside observer, as this was an opportunity to challenge unwarranted beliefs. Item 2 ("Exploration, curiosity, and a not-knowing stance") importantly tells the therapist to be active and curious. In Session D, there are many examples of the therapist taking a knowing stance, and such closing the door to training of mentalizing, and instead joining the patient in pretend mode (Table 10). The therapist here guesstimates instead of asking openly and joining the patient in the scene she describes. The end result is that the therapist presents an answer to the patient without first invoking any Socratic curiosity, asking for permission to present her own understanding. The therapist says, "This has been a… really an intense experience for you". To make this intervention mentalizing, she could have said: "From the sideline, it looks like this has been an intense experience for you". In this way, she would have made her own mind accessible to the patient, and the therapist would enter the scene with the patient and show/explain why she thinks like she does. Instead of stating that "What you are in the middle of now has to do with that rape", she could have described what she observes that makes her think that what the patient is in the middle of now has to do with the rape.

The therapist needs to be able to see the scene the patient paints with their words and then enter a conjoint process of trying to understand different perspectives in this scene while indirectly teaching strategies for understanding oneself from the outside and understanding others from the inside (mentalizing). Let us look at another such missed opportunity from Session D (Table 11). Here, the therapist could, for instance, have used Item 2 ("Exploration, curiosity, and a not-knowing stance") and simply asked what a new beginning means and explored the patient's thoughts on this subject. Paper II showed that perhaps the most important task for the MBT therapist is to maintain a flexible, playful mind that can mentalize well despite the patient demonstrating low levels of mentalizing. However, such interventions are not random. Next, we investigate how to navigate MBT sessions.

4.2.10 Item 12 and overall strategy for navigating and structuring sessions

One central feature of highly rated sessions of MBT is that the therapist never loses their overall goal (increase the patients' mentalizing) and remains steadfast and committed to this. Intervention number 12 ("Stop and rewind") would be one way of returning to any topic, but it can also shine through any other intervention whereby the therapist simply redirects attention to their chosen agenda. As the MBT-I manual does not contain any step-by-step recipe for how to structure sessions but is rather based on a fundamental understanding of the psychodynamic process in general, it is illuminating to see how this process plays out in

highly rated MBT. The first theme the patient introduces in Session A is that she cannot take in (believe/accept) compliments for having successfully arranged a funeral over the last six weeks. This is a melody the therapist recognizes and is therefore able to explore openly ("have you always had the role of a helper?" [Item 2]) to remind the patient this is a recurrent issue ("but what could this old way of behaving be all about?" [Item 2]), to define/confirm it ("yes, it feels really good for you to get that approval"), to validate her difficulties telling others about her success ("most people find that a bit difficult" [Item 13]), to challenge it ("why should others have yelled at you if you had not made it?" [Item 3]), and to ask her how it felt to finally get recognition [Item 11]. At 6 minutes into the session the therapist states: "Yes, you have been working on this for six weeks, you have succeeded, and you get credit for it". She then says: "It feels really good!" [Item 10]. As this feeling is now brought into the open and is part of the field between patient and therapist, the therapist then simply repeats this expression of healthy pride: "Yes, it feels really good?!" [Item 10]. Having achieved this aim, the therapist now focuses on the group.

Session C opens with the therapist not knowing if he has sent a statement on behalf of the patient. Instead of using this moment to initiate and display (exemplify mentalizing by, for instance, saying "I really want to help you with this statement, but now I am getting a bit insecure about what to say because I am actually not quite sure if it was sent. And I wonder how that will make you feel. I am actually a bit anxious about it, which may be connected to our last session." See Table 12. When the patient states that she was angry with the therapist, he misses a second opportunity to train in mentalizing. Here, he could have asked a question such as "How could I have understood or noticed that?" or "Thanks for letting me know, but I am not sure why you were angry with me; could you help me understand this better?" or basically any intervention aimed at elucidating the patient's inner workings regarding why she was angry with the therapist, why she tells him now, how this affects her, how she knows she is angry, etc. Instead, he enters a teleological stance where he heads directly into checking whether the statement was sent.

4.2.11 Item 17: The conjoint aspect of MBT
The therapist's focus on the conjoint aspect of MBT (Item 17) is crucial for beneficial treatment (Table 8) and reflects the alliance to the overall treatment program. In both sessions A and B, we see an intense focus on the conjoint aspect of MBT (Item 17 "Integrating experiences from concurrent group therapy"), combining MBT-I and MBT-G. This item,

which is by far the most frequent (mean = 49) in highly rated MBT, pertains to the overall program and is of course at the root of establishing a strong alliance. In low-rated MBT sessions, there are few interventions about the group (mean = 3). In the example above (Table 13), the therapist achieves both of these aims at the same time, as he does inquire about the patient's absence from the group in a transparent way. This makes his own mind accessible to the patient (Item 1), and she learns that he pays attention to her and that her actions have an impact on him (Items 14 and 15). Such strategic competence gave the therapists a broader roadmap of how to navigate, adjust, and tailor the MBT technique to the unique patient, relationship, and situation. This strategy is again rooted in the MBT theory, which in this case points towards epistemic trust and towards a recognition of how important it is for Elsa (Session A) to attend the group session (especially as there has recently been a rupture in the alliance with the group). A good working alliance makes the patient able to learn from the therapist at a significant level (necessitates that the patient feels understood). This alliance can be used to foster a good working alliance with the group, which is considered essential in MBT. To learn from the therapist, the patient first needs to be open to the therapist's knowledge about the inner and outer world. This ability is captured in the term epistemic trust (Fonagy et al., 2015a). By developing epistemic trust, the patient's ability to learn from the therapist (and consequently from others' minds in general) increases, and by learning mentalizing from the therapist the patient gradually learns how to learn on their own. Therefore, mentalizing is both a key to unlocking the patient's epistemic trust (its not-knowing stance applies for permission to enter the patient's inner chamber and turn up the light to increase understanding and clarity vis-à-vis the patient's inner processes) and the very process the patient needs to assimilate in order to improve (Fonagy et al., 2002). To be included and accepted by a group, one needs to master the implicit rules of conduct and manners. The more competent one is at this, the higher social rank one has the potential to achieve.

> In fact, being a competent member of a particular cultural or subcultural group means having at hand the implicit cultural meanings shared by members of that group. When a set of implicit cultural meanings has been shared over time, people do not have to refer explicitly to a particular meaning for it to be invoked. (Magnusson & Marecek, 2015, p. 143)

Patients with BPD are aware of this social game but are unable to master its rules. By chasing this enigmatic code, they lose sight of themselves and are mesmerized by the image they construct from the puzzle pieces of what they believe others need or want them to be. The

paradoxical result of misunderstanding the group norms is that BPD patients become even more bound by these rules or rather by their private version of these rules. BPD patients frequently have the same ingredients, but something does not quite add up for them. They often realize that their version of these norms is different than the consensus. When BPD patients are emotionally calm, they typically have an adequate ability to play the social game. This is one of many possible examples of how theory, knowledge about the patient, and the alliance work in concert to orchestrate such a display of strategic competence and battling. As BPD pathology often involves rigid thought processes. Skillful challenging of such ideas was found crucial in Paper II. Next, we turn our attention to the two MBT items especially designed for this purpose.

4.2.12 Items 3 and 9: Challenging maladaptive patterns of thinking

In session C, the therapist attempted several interventions targeting the patient's psychic equivalence, but he does not follow them up and may have failed in validating the patient's feelings before challenging them, which seemingly cements the non-mentalizing position. Importantly, in the low-rated MBT sessions Item 3 ("Challenging unwarranted beliefs") was absent, while in the good sessions, Item 3 averaged 6.5. In Session A, the patient's willingness to explore her own patterns and reactions is a result of an adequate working alliance carefully constructed over 14 months. Based on this solid ground, the therapist moves on to several "high-risk interventions" (Item 3: "Challenging unwarranted beliefs"; Item 9: "Psychic equivalence"; Item 14: "Transference and the relation to the therapist"; and Item 15: "Use of countertransference") during the session. Perhaps the most striking watershed between highly rated and low-rated MBT is the degree to which the curious, open, and not-knowing position becomes a vehicle for questioning, challenging, investigating, and identifying different perspectives in the patient's experiences. The more the relationship allows for contrasting views, confrontations, tolerating feelings, humor, self-disclosure, and curiosity, the more likely it is the patient will improve their mentalizing capacity. This session contains significant amounts of all these ingredients, and the therapy has evidently been important and helpful for the patient. As mentioned earlier, at the beginning of the session Elsa states "Oh, damn, what a skillful therapist you are! Thanks." As outlined above, session A revolves around three major themes: (i) false humility for having arranged a funeral, which turns into healthy pride, (ii) reluctance to be part of the group, and (iii) fear of not being able to shield her son from her own feelings (rather than providing comfort) as he threatened to commit suicide. The therapist arrests the patient's unquestioned beliefs across all three topics

and directs the mentalizing discourse to facilitate changes in perspective but not to present solutions or conclusions (as recommended by the manual; Karterud & Bateman, 2010). In the middle of the session, Elsa's therapist makes use of his own countertransference and challenges the patient's resistance against the group in a transparent way (see Table 13).

4.2.13 Items 16 and 2: Where pedagogy and curiousness intersect

It is important for us to note that although all the 17 items are distinct, some overlap considerably. Item 2 colors many other interventions in the MBT spectrum. Low-rated MBT seems to be characterized more by the use of Item 16 than Item 2, which then implies that the therapist follows and tries to understand the patient instead of being a not-knowing (curious) door to new knowledge. However, Item 16 can also be used in a more pedagogical manner, which has a different character we will return to in the Discussion section.

4.3 Paper III

Paper III demonstrated satisfactory levels of initial working alliance among BPD patients in MBT irrespective of clinical outcomes. MBT therapies with good outcome were characterized by a temporal increase in alliance strength as reported by the patients: In the model with Goals as dependent variable, the predictor indicating subgroups with good and poorer clinical outcomes accounted for 23% of the slope variation for this WAI-S/SR subscale. Corresponding models with bonds and tasks accounted for 25% and 35% slope variation, respectively. Comorbid paranoid PD was more frequent in the subgroup with poor outcomes, and also associated with poorer alliance development in this subgroup. However, there were patients with comorbid PD in the good outcome group as well (i.e., achieving an alliance with these patients is difficult, but not impossible). Differences in alliance development according to outcome were most pronounced for the subscale tasks.

4.3.1 Descriptive data in subgroups with different outcomes

Patient factors

The patients' baseline levels indicated severe problems with functioning and distress at treatment onset but no significant differences in severity by outcome subgroup. The sample was characterized by , patients with BPD, reports indicating poor QoL, considerable comorbidity, and personality problems across all domains, especially within the domains of identity and self-control. The good outcome subgroup was characterized by younger age,

fewer patients with no months of work/study at all previous year, and fewer with comorbid Paranoid PD and mood disorder. In the preliminary analyses, age and paranoid PD explained 2%–5% of the variation in GAF slope.

Treatment factors

Nearly all patients in the good outcome subgroup completed treatment according to plan (91%), versus 58% in the poor outcome subgroup. Mean treatment duration was 27 months (SD 13), early drop out (< 6-month duration) was minimal (2.5%), and did not differ by subgroup. In the good outcome group, there were no later drop-outs, while 9% were later drop-outs in the poor outcome group.

4.3.2 Main analyses: Longitudinal course of working alliance

Initial levels of working alliance (all subscales) were well within an acceptable range, and there was a significant increase of all three working alliance subscales over time. There was significant longitudinal between-subject variation. These change patterns also remained significant in models a) controlling for variation associated with different WAI versions and b) investigating possible bias of different assessment numbers.

Variation associated with good and poor outcome subgroups

The good and poor outcome subgroup predictor was investigated in each of the three models. Initial levels of Goals, Bonds, and Tasks did not differ by subgroup, but change over time was significantly different by subgroup. The subscale Goals accounted for 23% of the WAI-S/SR slope variation, Bonds for 25%, and Tasks for 35%. These findings remained significant for the three subscales—Goals, Bonds, and Tasks—in models a) controlling for variation associated with different WAI versions, b) investigating possible bias of different assessment numbers, and c) corresponding differences were also found in models investigating the dichotomous WSAS and BSI outcome variables as predictors. In the good outcome subgroup, ratings of Goals, Bonds, and Tasks increased significantly over time. In the poor outcome subgroup, change over time was not significant for any of the WAI-S/SR subscales.

Variation associated with patient factors

Relevant patient factors (age, comorbid mood disorder, and comorbid paranoid PD) were investigated as separate predictors added to the three WAI-S/SR subscale models. Mood disorder was associated with significantly lower initial alliance levels, but not deviating

change over time. Age was not associated with significantly deviating initial alliance levels or deviating change over time, but explained some longitudinal variation. Paranoid PD was not associated with baseline deviation of WAI-S/SR ratings in any of the two outcome subgroups. The presence of paranoid PD was associated with poorer development of WAI-S/SR subscales over time in the poor outcome subgroup, but not in the good outcome subgroup. Corresponding results for paranoid PD were also found in models investigating the supplementary dichotomous WSAS and BSI outcome variables. In the good outcome subgroup, ratings for goals, bonds, and tasks increased significantly over time (for all $p < 0.05$). In the poor outcome subgroup, change over time was not significant for any of the WAI-S/SR subscales ($p > 0.1$).

5 Discussion

The overall aim for this dissertation was to investigate (e.g., identify, measure and differentiate) MBT. The approach was to identify missing parts in an ideal measurement of MBT, like searching for missing pieces of a jigsaw. We should remind ourselves and the reader that this thesis consists of three articles with the following aims: 1) to investigate the reliability of the MBT-G-AQS (Karterud, 2015); 2) to explore what characterizes MBT with high and low ratings in terms of common factors; and 3) to examine how the three facets of the working alliance developed in MBT with good and poor outcomes. The overarching aim of the thesis as a whole is to discuss how MBT can be integrated within the larger field of psychotherapy research. In this way, the thesis seeks to encourage a marriage of the CF and EST approaches by investigating one specific evidence-based treatment for BPD through the lens of the CFs. This dissertation is but a small atom in a vast molecular structure or in our case four pieces in a larger jigsaw puzzle. However, as we learn from the hermeneutic tradition, if one can be precise enough (e.g., describe in detail or "split an atom"), one can produce potent chain reactions and large-scale impact. In the present study, an overarching aim is to *sublate* (Hegel, 2018) the dialectic schism between the common and the specific psychotherapeutic factors in our attempt to measure MBT. After a brief recap of what motivated the study, the three articles will be discussed separately and then looked at as a whole. Subsequently, the marriage of CF and EST will be considered. This is followed by a description of the limitations and a discussion of the possible implications, such as arguing for an understanding of psychotherapy based on "embedded alliance" and the viewpoint that the different evidence-based treatments for BPD should work more in cooperation than is currently the case (as it seems they share core CFs and may hold some key pieces in each tradition that combined would deliver the best potion). Further, the indicated centrality of focus on tasks and goals (Papers II and III) signals the importance of pedagogy (e.g., epistemic trust), which will be elaborated and proposed as a fertile path for future studies. Paper I found that overall adherence and quality for MBT-G can be rated reliably, which has implications for both quality control and the reporting of fidelity measures for MBT. However, in line with the critique by, for instance, Sharp et al. (2020), the core theoretical components of MBT are not well operationalized, and it is an undermining finding that the reliability of the core concepts "Pretend mode" and "Psychic equivalence" is low. Paper II indicated that applying the "ongoing dialectic between technique, theory, and alliance" (e.g., CF & EST; Hatcher, 2010, p. 18) appeared to be fruitful in studying a specific evidence-based

treatment for BPD. Paper III showed that a positive development in the working alliance, particularly for the tasks and goals subscales, characterized MBT treatments with good outcome. This implies that an investigation into how the therapist tailors the technique to the unique patient (Paper II) was indeed called for and also provides direction as to where MBT research could focus next.

5.1 The importance of evidence-based treatments for borderline personality disorder

Within clinical psychology, the development of evidence-based treatments for psychiatric, social, emotional, and behavioral problems has been remarkable (Kazdin, 2011). This is true for PDs in particular. PDs have "traditionally been viewed as chronic disorders" (Kvarstein, 2013, p. 6), and the impact on society and individuals can be severe. In a Norwegian population study (Cramer et al., 2006) the general impact of PD on QoL was stronger than the influence of sociodemographic factors, physical health factors, and Axis-I disorders. Therefore, the emergence of evidence-based treatments for PDs has great implications for patients, their families, and society at large. It also has implications for therapists in terms of being able to rely on a larger program when dealing with patients with PDs. As will be elaborated below, it has political implications (Duncan et al., 2010), consequences for epistemic trust, inferences in terms of promoting therapist allegiance, and very likely a placebo effect (Wampold et al., 2007b).

5.2 Fidelity measures for mentalization-based treatment

This thesis argues that competition between different evidence-based treatments is futile (e.g., Messer & Wampold, 2002) and that increased collaboration can create a better understanding of how these treatments work and in time perhaps also what works for whom (Antonsen, 2016; Roth & Fonagy, 2006). "Instead of understanding how our treatments work, we somewhat mindlessly repeat exactly the behaviours that led to the positive observed outcomes" (Lemma et al., 2011, p. 16). However, before trying to understand why something works, we should perhaps at least try to identify what it is (i.e., to operationalize and define it). An important but typically neglected necessity in this regard is reliable and valid fidelity measures. It has been suggested that by using appropriate assessment measures, it is possible to measure the active ingredients of change within the different therapeutic models and possibly shed some light on the Dodo bird verdict (Shedler, 2010). Measures of treatment integrity are also crucial in order to claim we have an evidence-based treatment. Within the realm of evidence-based treatments for BPD, we still see few studies presenting what seem

like valid and reliable instruments to accompany a treatment's superior status. As we have seen, fidelity was poorly reported in almost half of the studies (Malda-Castillo et al., 2019), which is likely connected to the required resources involved in developing and implementing integrity instruments. This number is surprisingly high; for example, in a review of randomized psychotherapy studies, Perepletchikova and co-workers (2007) found that only 4% of studies documented treatment integrity. It seems that Malda-Castillo et al. (2019) may have been rather generous in their criterion for what accounts as a valid fidelity measure. This is also the case in the previously mentioned study by Kivity et al. (2019), who measured adherence in a study of evidence-based treatments for BPD using the PQS. A justification of such reasoning requires a closer look at this scale, which will indicate the importance of developing proper psychotherapy measures, especially when applied to the evidence base of specific treatments.

Notably, the (PQS; Jones, 1985) is a useful instrument and as will be suggested below should probably be included in the toolbox used in the Quality Lab for Psychotherapy (especially as we have been certified PQS raters since 2009). As Kazdin (2009) noted, to understand how psychotherapy produces benefits researchers need to investigate the interplay among processes. However, despite both the working alliance and other CFs being "quintessential" change agents (Wolfe & Goldfried, 1988), we should still try to measure/identify the specific ingredients in each psychotherapeutic potion carefully. Regarding the PQS, Jones (1985) states that while "built on general assumptions of psychotherapy as an interpersonal process, it is intended to be neutral with respect to any particular theory of therapy, and should permit the portrayal of a wide range of therapeutic interactions" (p. 2). I find it somewhat scientifically suspicious that an instrument "intended to be neutral with respect to any particular theory of therapy" (ibid.) is being applied as a measure of treatment integrity. As good atomists, let us be even more precise. The PQS consists of 100 CFs weighted from 1–9 by the rater. Therefore, this instrument will characterize the treatment, or more precisely the therapist's activities, as involving more or less of different CFs. Consequently, one would expect a dynamic psychotherapy, where different CFs will be flexibly adopted in each session, to have a wide variety of representations (ratings) in terms of the PQS (reliable but not always valid). In a recent review to summarize the major evidence-based psychotherapies for BPD, (Ellison, 2020) states that "the low therapist competence ratings in the TFP arm raise questions as to how faithfully TFP was delivered in this trial—a noteworthy limitation, as conformity to TFP technique is an important predictor of outcome", a statement based on

the study by Kivity et al. (2019). We find it a "noteworthy limitation" to predict outcome for an evidence-based treatment based on the PQS, as the very purpose of the 100 items of the PQS is to provide a basic language to describe and classify therapy processes. Let us consider several PQS items to make this point sufficiently clear. Item 4 is "The patient's treatment goals are discussed", Item 6 is "Therapist is sensitive to the patient's feelings, attuned to the patient; empathic", Item 32 is "Patient achieves a new understanding or insight", and Item 46 is "Therapist communicates with patient in a clear, coherent style". The description of Item 46 is "Place toward characteristic end if therapist's language is unambiguous, direct, and readily comprehensible. Rate as very characteristic if therapist's verbal style is evocative and marked by a freshness of words and phrasing. Place toward uncharacteristic end if therapist's language is diffuse, overly abstract, jargon-laden, or stereotypic". Consequently, as an instrument to characterize sessions in terms of their focus on different CFs, PQS is highly useful (and could be used in tandem with specific measures). However, it is possible to argue that PQS is better at describing the trajectory of each therapy than a valid measure of adherence to the treatment. It would be walking on thin ice to induce that because something is wet, it must be water; even water, it appears, does not need to be wet. As Kivity et al. (2019) is the only identified study that actually reports integrity measures for comparisons of different evidence-based treatments for BPD (although based on PQS), studies like Paper I of this dissertation seem to be needed.

5.3 Paper I

Let us first address the topics and four major findings highlighted by Paper I:

1) The group component is typically neglected, both within the field of MBT and in the larger universe of psychotherapy research.
2) The overall score for adherence and quality of MBT-G-AQS can be reliably rated by one rater.
3) The items measuring theoretical constructs considered core concepts in MBT showed low reliability, both "Pretend mode" (Item 15) and "Psychic equivalence" (Item 16).
4) The nine group-specific items displayed high reliability for both adherence and quality.

The first topic that formed the background for Paper I was the paucity of research on adherence and competence for group therapy, confirming the call for the development of

group therapists' measures (Burlingame et al., 2004). The Group Psychotherapy Intervention Rating Scale (GPIRS) was developed by Sternberg and Trijsburg (Chapman et al., 2010) and is the only scale that seems to reflect the many integrity measures developed for individual psychotherapy. GPIRS was developed for group psychotherapy in general. The 48 items are designed for empirical research in general and are not specific for any treatment or manual. The MBT-G-AQS addresses the dialectic between structure and dynamic process, which is present in all dynamic group therapies (Yalom & Leszcz, 2005). Therefore, Paper I and the scale itself should be of interest for the general field of group psychotherapy, and the MBT-G-AQS should be helpful for the future development of other similar scales for other group psychotherapies.

Treatment integrity consists of two elements: (1) treatment adherence, that is, "the extent to which a therapist used interventions and approaches prescribed by the treatment manual, and avoided the use of interventions and procedures proscribed by the manual (Waltz et al., 1993, p. 620) and (2) the therapist's competence, that is, "the level of skill shown by the therapist in delivering the treatment. By skill, we mean the extent to which the therapist conducting the interventions took the relevant aspects of the therapeutic context into account and responded to these contextual variables appropriately" (Waltz et al., 1993, p. 620). According to this definition, competence requires adherence, but adherence does not necessarily imply competence (McGlinchey & Dobson, 2003). In RCTs "where therapists are trained using a manual for a specific disorder, between-therapist variation is likely smaller than in general practice" (Falkenström et al., 2013b, p. 2). Therefore, despite the "robust" therapist effect reported (Wampold & Imel, 2015), a reliable integrity measure based on a manual has important implications for delivering a specific potion to BPD patients. A manual is also important in making therapists, and indirectly their patients, trust their method (therapist allegiance). The group component is the clinical backbone of the MBT program (Karterud, 2015). However, the group component had been neglected when it comes to fidelity measures. This is unfortunate, as a reliable fidelity measure is important not only for reporting treatment integrity but also for quality control, supervision, training of therapists, legitimization of the treatment (e.g., government, propagation), and further research and proliferation.

The second major finding in Paper I was that the scale showed high reliability. The present reliability is somewhat higher than that reflected by the G-coefficients in the reliability study

on the MBT-I-ACS (Karterud et al., 2013), a possible product of extensive training and experience. With one rater, the reliability was very high for overall MBT-G adherence (.86) and quality (.83). This demonstrated that the MBT-G-ACS can be reliably used by one rater to determine the cut-off for adequate adherence and quality/competence for MBT-G. The reliability for overall adherence and quality ratings with five raters were high. This indicates that a team of raters was able to achieve good agreement regarding the ingredients in MBT-G are and how to evaluate them. The reliability for the overall absolute decision (absolute G-coefficients) was very good. As items with low absolute G-coefficients also had low variance, the reliability is therefore deemed high for these items as well (Hagtvet, 1997). The scale may contribute to future psychotherapy research by assuring internal validity and contribute to research on adherence and competence as possible moderators of treatment outcome. In addition, the scale can be used for training and clinical purposes; assessing and providing feedback about therapeutic competence and adherence enables therapists and supervisors to check and improve the skills used in delivering essential elements of MBT-G. Noticeable differences in the mean profiles for MBT-G and PDG are interpreted as reflecting the scale's ability to differentiate these two treatments, thus lending some support to the discriminant validity of the scale. The results were both uplifting in terms of demonstrating that the overall score for adherence and quality (competence) could be rated by one single rater and that the overall scale has good reliability for two raters. This finding is important in terms of the feasibility of integrating quality control and assessment at multiple treatment facilities and of continuing the MBT-G ratings made for services such as the Quality Lab for Psychotherapy in Oslo. The good inter-rater reliability results in this study indicate that the MBT scales can be used reliably with careful training and supervision. Nonetheless, subsequent studies should investigate whether this finding can be replicated with other raters. A limitation of reproducibility (which is at core of reliability) is whether such agreement can be reached at other places/centers and whether the MBT-G-AQS is primarily a tool for expert raters with special training (Simonsen et al., 2019). Recently, the MBT-G-AQS was employed to measure treatment fidelity in a Danish RCT (Beck et al., 2020; Jørgensen et al., 2020) and has been reported in two studies by Kvarstein et al. (2019, 2020).

The overall ratings in Paper I were based on a global understanding of the session, which is essentially about answering whether the therapists stimulate the patients mentalizing or not. That is, the "most important sign of a successful MBT session is that the patient gets involved in a mentalizing discourse" (Karterud & Bateman, 2010, p. 44). This raises another concern,

which is that it is not possible to rate the therapist(s) independently of the patient(s). It has generally been assumed that adherence and competence are therapist characteristics (Baldwin & Imel, 2013). Recent studies (e.g., Boswell et al., 2013) challenge this presumption, and it seems that "it is the patient's contribution to competence ratings that is related to outcome rather than the therapists' competence relative to other therapists" (Wampold & Imel, 2015, p. 238). No matter how well defined the scale and manual, it will be necessary to rate the interaction between therapists and patients. This implies that a substantial portion of the variance in both adherence and competence ratings will stem from the patients. Further, the conception of competence/quality in MBT should thus be derived from the treatment manual and the theory of change specified in it. However, MBT is a manualization of a non-technique-based psychotherapy (Perepletchikova, 2007; Perepletchikova et al., 2007; interventions are driven by understanding), in which the relationship to therapist and interactional processes play an essential role. Consequently, as indicated by Paper II, highly rated MBT contains strategies not described in the manual, which may imply that the concept of quality in MBT is largely a measure of embedded alliance (this topic will be elaborated later). Paper II investigated such conceptual interactions closely in MBT-I, but case studies in MBT-G should also be applied to investigate this topic further.

The third major finding was that some of the items measuring core MBT concepts had low reliability and occurrence (e.g., "Psychic equivalence" and "Pretend mode"). An important aspect of a reliability study is identifying items in the manual that should be made more precise. For example, items with the lowest reliabilities in MBT-I following a brief 1-day training course were "Focus on affects", "Focus on interpersonal affects", "Counter-transference", and "Psychic equivalence" (Simonsen et al., 2019). For MBT-G, this is particularly true for psychic equivalence and pretend mode (these were the two concepts raters disagreed about the most), which is somewhat unfortunate, as they are part of the core theoretical underpinning in this treatment tradition. The G-study allows for investigating the source of variance, and for these two items the results indicate that the measured concepts are unclear for therapists and raters alike. This finding is largely in line with Karterud et al. (2013)'s finding regarding the MBT-I-ACS that "there was a moderate agreement on identifying interventions aimed at psychic equivalence. However, the competence reliability is lower (.33). The manual should be more specific with respect to what counts as a high versus low competence for this item" (p. 714–715). In terms of pretend mode, they reported that

the residual variance was very high for this item, indicating (1) that the therapists had difficulties with identifying pretend mode, (2) that the therapists had difficulties with knowing what to do with it, and (3) that the raters had difficulties with identifying interventions aimed to modify pretend mode. (p. 714)

Hence, Paper I and the studies by Simonsen et al. (2019) and Karterud et al. (2013) may lend support to criticism aimed at MBT being abstract and hard to integrate (Hutsebaut et al., 2012; Sharp et al., 2020). Consequently, amending the operationalization of pretend mode and psychic equivalence will most likely be helpful for the field of MBT. From a psychometric perspective, items/interventions with low occurrence (e.g., Items 9, 14, 16, and 18) may be seen as redundant. Further, very high reliabilities (.95 or higher) are not necessarily desirable. Such views also reflect the underlying question of whether MBT-G is best defined by empirical data (what can be operationally observed in therapists who say they deliver MBT-G) or by an a priori conception by the conceivers of MBT. Arguably, there is an interaction/dialectic between such perspectives with clinical practice, such that over time there will be an interplay leading to a continual revision of manuals, theory, training, rating procedures, and practice.

It is interesting to note that in Paper I there was a difference between treatments (PDG versus MBT-G) on competence but not adherence, especially as much of the previous research (Barber et al., 2007; Gutermann et al., 2015) has reported that the interrelatedness of adherence and competence is high. One possible reason for this finding is that mentalizing is a very broad concept (CF; "Plain Old Therapy", e.g., Allen, 2012). Therefore, most treatments would necessarily deliver mentalizing interventions but with different competence/quality. Further, the structuring aspect of the treatment is assumed important in a clinical setting (Bateman & Fonagy, 2016; Wampold & Imel, 2015). Presented findings would support this, as it was the structuring elements of the MBT-G-AQS (Items 1, 2, and 3) of MBT-G that constituted the major difference from PDG. Inderhaug and Karterud (2015) reported that without this structuring element, MBT-G groups can be very chaotic.

The fourth major finding in Paper I was that the nine group-specific items displayed high reliability for both adherence (range .83–.95) and quality (range .78–.96). This means that the operationalization of MBT-G (Karterud, 2015) has been fruitful. The combination of group and individual therapy (conjoint therapy) has been found to be positively associated with outcome (Antonsen et al., 2017). Before discussing Paper II in more depth, it is telling to

observe that the highly rated MBT sessions in Paper II displayed an impeccable focus on the conjoint aspect of MBT (Item 17; "Integrating experiences from concurrent group therapy"), that is, combining MBT-I and MBT-G. This item, which is by far the most frequent in these two sessions, has to do with the overall program and is of course at the root of establishing a strong alliance in the overall program. This item is important because it builds a bridge between the individual and their place in society (the group is a small society or "family"). As "personality disorders are defined as different ways of organizing social experience" (Pedersen, 2008, p. 72), this is likely one of the main keys BPD patients need to improve. In the low-rated MBT sessions, there were few interventions about the group. The importance of the conjoint aspect of MBT will be discussed when covering pedagogic interventions and epistemic trust in more depth.

5.4 Paper II

The following four major findings from Paper II will be discussed in more detail below:

1) The variation between highly rated and low-rated MBT can be investigated in terms of CF concepts.
2) Highly rated MBT was characterized by a carefrontational style, where therapists battled the patients' comfort zones in a tailored fashion, while displaying faith in their own method (e.g., Falkenström et al., 2013b, p. 10).
3) Avoid being overwhelmed by countertransferences.
4) Reintroduce "embedded alliance".

Individual MBT already had reliable fidelity measures for both adherence and competence (Karterud et al., 2013). The Quality Lab for Psychotherapy at Oslo University Hospital applied this scale and rated 108 individual sessions over a 3-year period, with at least two raters per session. Having established and implemented a reliable scale, our next question was what characterized sessions with different levels of ratings. However, the answer to what characterizes good MBT could not simply be that it was "good MBT", that is, displaying many MBT interventions of high quality. Such an article would provide a list of effective strategies in MBT or goods examples of MBT interventions and would be valuable in clinical practice but would not expand our understanding much. As the answer in this case is (part of) its own definition, we would expect such a path to process research, similar to mathematics trying to explain itself through mathematics: "no more than a mathematician can show by way

of mathematics – by means of his science, that is, and ultimately by mathematical formulae – what mathematics is" (Heidegger, 1976, p. 33). That is, such tautological logic would be empty of explanatory power. Therefore, when contemplating the arena of current psychotherapy research, the domain of PDs is one of the few demonstrating the superiority of certain treatment programs/methods despite the Dodo bird verdict being alive in terms of no reported difference between the evidence-based treatments for BPD (Ellison, 2020). Our main question was how therapists in an evidence-based treatment for patients with relational pathology and low epistemic trust (BPD) tailored their technique to the unique patient. Therapists seeking greater efficiency learn that for unknown reasons some therapists excel in manifesting strong working alliances, no matter what method they employ (Lemma et al., 2011). Falkenström et al. (2013a) reinforced the conclusions of Baldwin et al. (2007), which showed that only the mean level of alliance for each therapist was important for outcome. "The within-patient effect of alliance on symptom level varied significantly between patients, but not between therapists" (p. 326). Therefore, we saw a need to investigate how such skilled therapists fostered the therapeutic alliance. Manualized treatments can be viewed as attempts to provide aspiring clinicians with some guidelines from expert therapists on what kind of strategies or interventions are considered to be helpful or to nurture the alliance for certain problems or types of patients (e.g., Lemma et al., 2011). However, the manual (treatment approach) must be adapted to the specific patient, and this merger of the working alliance and the specific technique(s) was the topic of our IPA.

One of the first observations in our IPA was that highly rated MBT, not only includes, but transcends the manual. These therapists seemed to possess some sort of inner map, knowledge, or understanding of the patient and the present situation that made them navigate the interpersonal terrain in a flexible way without losing track of their long-term goal(s). The interventions built logically on each other and seemed guided by an overarching strategy: "If one intervention failed, the therapists pursued the same goal by another route. In the low-rated sessions, interventions were infrequent, and often lacked a clearly detectable plan or overarching pattern" (Paper II). One central question then emerges: What does skillful MBT look like without primarily using the language of MBT? Can the CFs explain the difference between highly rated and poorly rated sessions? Our results indicated that one can indeed explain MBT in terms of the CFs but also that the therapy was focused on an increase in mentalization. The overarching strategy or strategic competence was closely linked to the working alliance and epistemic trust. It could be seen as the best attempt to address the

challenge: Given the patient, goal, situation, and relationship, how do we best bring about change? Strategic competence may provide the therapists with a broader roadmap of how to navigate, adjust, and tailor the MBT technique to the unique patient, relationship, and situation. Strategic competence partially overlaps with the quality score of MBT; it includes the timing, precision, and relevance of the interventions. Consequently, a skillful application of MBT includes an overarching ability to navigate without being defined by the MBT manuals. Therefore, one problem with manuals (attempts to transmit knowledge from expert therapists) for the average therapist is that they do not teach such overarching strategies, something MBT has been criticized for in terms of being difficult to learn and operationalize (Hutsebaut et al., 2012; Sharp et al., 2020).

The first of the four major findings in Paper II was that we seemed able to denote differences between highly rated and poorly rated MBT in terms of CF concepts (the observed variation between highly rated and poorly rated MBT could be investigated in terms of CFs). Highly rated MBT contained overarching strategies to systematically challenge the patients' world view that are not defined by the manual. As the alliance can be seen as looking at the relationship in terms of meaningful work (Hatcher, 2010, p. 25), this may signal that the rated competence/quality of MBT was associated with alliance strength and that the observed quality of a treatment is some sort of measure of the "embedded alliance" (Hatcher, 2010). As we have seen, in Bordin's view (1979) different types of psychotherapy need different types of alliances. For example,

> a treatment geared toward changing deep personality structures would depend much more on a strong emotional bond between therapist and patient for the patient to feel secure enough to engage in the emotionally painful therapy work than exposure therapy for a simple phobia, which probably depends more on agreement on tasks. (Falkenström & Larsson, 2017, p. 167).

It seems likely that every type of therapy will promote/foster somewhat different alliances (Bordin, 1979). Bordin also highlights the importance of the bond when working with difficult material: "Some basic level of trust surely mark all varieties of therapeutic relationships, but when attention is directed toward the more protected recesses of inner experience, deeper bonds of trust and attachment are required and developed" (Bordin, 1979, p. 254). With BPD patients, different techniques are demonstrated to produce different alliances; for example, the results reported by Spinhoven et al. (2007) "indicate that the rating

of the alliance reflecting the overall quality of experiences and feelings during a large number of therapy sessions clearly differs between treatment conditions" (p. 112). The battling style found in highly rated MBT in Paper II indicates that a focus on tasks and goals may be particularly important in treatments of BPD. This is not in disagreement with the point made by (Falkenström & Larsson, 2017) but could simply indicate that to foster a strong bond, one would need to focus strictly on the tasks and goals of therapy and that the bond part is strongly connected to (patients') epistemic trust acquired from previous challenges in terms of a mentalizing discourse. The findings in Paper II are in line with such reasoning. Such a process seemed to create a positive feedback loop between the bond (relationship) and the ability to focus on tasks and goals. Paper II also indicated that the bond was an important asset for the therapist to be able to challenge adequately by focusing on the tasks and goals of therapy. In fact, this seems to be in line with Falkenström et al. (2015), who suggest that tasks and goals are one factor and not two. Wampold and Imel (2015) highlight the importance of a healing ritual, agreement on the explanation of the problems, and a cure congruent with this conceptualization. This indicates the need for a strong agreement and focus on tasks and goals in therapy. Paper II signaled that maintaining a positive personal bond seemed less important than using that bond for meaningful work, such that the bond part of alliance in the low-rated sessions was also partly positive but lacked meaning, purpose, and direction. Epistemic trust seemed like the asset that was built from alliance ruptures and repairs and adequate battles of the comfort zone. Therefore, the current investigation of MBT highlights that therapy is not about simply being supportive, having a "tea party", avoiding difficulties, or being uninvolved in the relationship. Importantly, Paper II proposes that the part of the bond that grows in time seems to be epistemic trust, which seemed associated with a strong focus on tasks and goals in MBT. The development of embedded alliance measures may prove pivotal for the further integration of the common and specific factors in psychotherapy research, which is likely necessary for the increased understanding and measurement of specialized treatments (e.g., MBT). We will return to this later in the discussion.

The second finding worthy of special attention was that highly rated MBT was characterized by a carefrontational style, where therapists battled the patients' comfort zones in a tailored fashion. This is not only in line with the importance of "psychotherapists' deep engagement in the client's welfare, willingness and capacity to confront the client's dysfunctional behavior, maintenance of optimism and a resource-focus while also being playful" (Rønnestad, 2016, p. 12; Råbu et al., 2011) but also underscores what Fonagy et al. (2019) denote the "re-

emergence of social learning", which they claim is "the way in which any effective treatment is embedded in metacognitive processes about the self in relation to perceptual social reality" (p. 94). The therapeutic relationship should enable the patient to develop other learning relationships based on an acquired sense of how to trust another person as a source of significant social information. Consequently, one central outcome from studying MBT in detail is that a steady mentalizing approach seems like a close "technical approximation" of optimal reparenting; we see more clearly how the specific technique influences alliance building and epistemic trust. It is possible to argue that to achieve an attachment with a BPD patient (alliance) who will largely be healed by that corrective experience itself (Fonagy & Bateman, 2006), epistemic trust can be attained by attending to and caring so deeply for the patient that the therapist challenges the patient's deepest belief systems in a way that makes it safe to trust a new way of experiencing reality. In treating BPD, "[m]uch of the therapist's role consists of a process of reeducation, and in the course of time the therapist even becomes a role model for the patient" (Spinhoven et al., 2007, p. 104). In line with this, Høglend (2014) has shown that working with the countertransference—that is, addressing the therapeutic relationship—is an effective strategy with BPD patients. This could be seen as interpersonal psychopedagogics. Paper II demonstrates the importance of tasks and goals to achieve a strong bond (epistemic trust) when working with severe pathology and the importance of a pedagogic stance (building epistemic trust). This makes the re-educational or reparenting aspect of (BPD) treatments (Spinhoven et al., 2007) more apparent. As we will discuss later, in terms of Kierkegaard's (1998) idea of helpful relations, such tailored battles seem to necessitate that the therapist listens deeply to the content and nature of the narrative (e.g., narrative identity; Lind et al., 2019a; Lind et al., 2019b), implied worldview, and mentalizing process. In other words, the therapist must strongly empathize but not identify with the patients' narrative, while being able to challenge them and alter the discourse. For this reason and because a "pedagogic stance" has recently been prescribed by the MBT manual (Karterud et al., 2020), the thesis will return to a more in-depth examination of pedagogics in MBT.

Importantly, one "of the sacrosanct assumptions of a client is that their therapist believes in the treatment being delivered" (Falkenström et al., 2013b, p. 10; Wampold & Imel, 2015, p. 120). This is a topic that will be discussed in depth later, for example, when the placebo effect is addressed. The second major finding in Paper II, coinciding with the CF approach, is that the therapist staying on course and battling the comfort zones of the patient is a way to

communicate trust in their own treatment method and thus instill hope. This again fostered epistemic trust and increased alliance. The literature supports this view and typically reports that the very administration of the "magic potion" may be as important as the active ingredients. Kaptchuk et al. (2008) reported that factors contributing to the placebo effect and non-specific effects can produce both statistically and clinically significant outcomes and that the patient–practitioner relationship is the most robust component. The authors concluded that "warmth, empathy, duration of interaction, and the communication of positive expectation might indeed significantly affect clinical outcome" (p. 7). The positive expectations instilled in the highly rated sessions were connected to the tasks and goals of therapy and seemed particularly important in MBT; this is further investigated in Paper III. The bond was an asset that the therapists could use to advocate the importance of the tasks and goals, and it also seemed crucial that the challenges—in line with the theory of an unconscious alliance by (Davanloo, 1990a, 1990b)—fostered epistemic trust. Such a theory would also highlight the importance of focusing on tasks and goals and the development of epistemic trust.

Paper II found that the bond part of the alliance in MBT seemed like an asset that could be used to promote focus on goals and tasks. The therapeutic process was some sort of battle, and in the low-rated sessions the therapists certainly were brought out of their comfort zone, to the degree that they abandoned the therapeutic project (temporarily). Luborsky (1976) brilliantly employed a counting signs method of assessing alliance and described two types of alliance, one "based on the patient's experiencing the therapist as supportive and helpful" and one "based on a sense of working together in a joint struggle" (p. 94); the "joint struggle" inspired us to title Paper II "Battles of the comfort zone". As Morken et al. (2019) state, "After all, therapy is not supposed to be a tea party" (Morken et al., 2019, p. 11). In Paper II, higher quality implied more battles, and therapists seemed to nurture the alliance through battles of the comfort zone.

In BPD treatments, "a sudden shift from idealizing to derogating the therapist can disrupt the patient's capacity to work with therapist comments and may result in unilateral termination on the part of the patient" (Levy et al., 2010, p. 414). Consequently, dealing with alliance ruptures becomes crucial in effective BPD treatments (Morken et al., 2019). In line with the findings presented in Paper II, Boswell et al. (2013)'s study on CBT found that higher levels of interpersonal aggression were associated with lower adherence and competence ratings and that both adherence and competence ratings deteriorated over the course of treatment. Anger

and hostility are widely recognized as difficult emotions to work with in therapy (Mayne & Ambrose, 1999) and are central features of BPD patients. As emphasized by Sharp et al. (2020), effective MBT presupposes an ability to not be overly involved in the content of the patients' narrative but rather to focus on the (mentalizing) process itself, something that is not easy, as borderline pathology is typically characterized by intense emotions and a tendency to trigger tough transferences (Colli et al., 2014). BPD "symptoms frequently interfere with the development of the therapeutic alliance and make treatment a long and difficult endeavor, fraught with recurrent ruptures, perceived empathic failures, chronic evasiveness, angry outbursts, and premature termination" (Levy et al., 2010, p. 413). As the MBT therapist needs to be an attachment figure for the patient (Karterud & Bateman, 2010), the ability to maintain high RF and relationally navigate the relational landscape in a manner allowing the patient to gain corrective emotional experiences will most often include patients testing whether the therapist can be trusted. Resulting battles of the comfort zone may (temporarily) weaken the alliance. According to Safran and Muran (2000), such alliance ruptures can be seen as a royal road to identifying and addressing the transference/countertransference. The results in Paper II indicated that high levels of epistemic trust seemed to be a product of previous alliance (rupture) processes. Observing these findings, one could wonder whether epistemic trust is a crucial factor for therapeutic relationships to foster healing effects. Relational healing is hard to imagine in the absence of epistemic trust. Consequently, as epistemic trust is considered lacking in BPD patients (Fonagy et al., 2015a), it would resonate with MBT theory if acquired epistemic trust could be of special importance for borderline patients, often considered the very prototype of personality pathology (Kernberg & Caligor, 2005; Sharp et al., 2015), that is, relational pathology, and trust issues, at core.

The third major finding in Paper II was to avoid being outplayed by one's own countertransferences. Another related relational reeducation (i.e., pedagogic stance) aspect is that the therapist needs to both avoid being handicapped by countertransference(s) and simultaneously allow for relational growth by addressing ruptures in the alliance (building a strong emotional bond and gaining epistemic trust). Therefore, one would expect that an effective therapist would need to adequately mentalize the patient's epistemic trust and alliance, allowing for such relationship building. One imaginative method to indicate therapists' ability to mentalize (i.e., RF) might be whether or rather to what degree they are able to identify and evaluate the therapeutic alliance as experienced by the patient. Therefore, despite therapist- and patient-rated alliances being equally good predictors of outcome

(Flückiger et al., 2018), it would be interesting to investigate whether therapies with good outcome are characterized by high congruence between patient-rated and therapist-rated alliance. Cologon et al. (2017) report that therapists' RF predicted therapist effectiveness. In their study, "secure attachment compensated somewhat for low reflective functioning and high reflective functioning compensated for insecure attachment" (p. 614).

Being aware of the different countertransferences typical for meetings with BPD patients should be facilitated by examples and case studies, preferably in the manuals but also from theory. According to Masterson (1988), the "borderline patient defines love as a relationship with a partner who will offer approval and support for regressive behavior" (p. 110).

Transactional models inform us that individuals impact their environment such that the characteristics of both the person and the environment change in ways that will alter the relationship (e.g., the nature of future interactions) between the two (Cicchetti & Rogosch, 2002; Steinberg & Avenevoli, 2000). For example, "disorganisation of the attachment system may cause a child to be increasingly manipulative and controlling over their environment, but such controlling actions may undermine the caregiver's capacity to provide a normative playful environment to his or her toddler" (Fonagy & Bateman, 2007, p. 84). The resulting lack of epistemic trust (Luyten et al., 2020b) needs to be addressed in BPD treatments, something that can be done using the three different communication systems proposed by Bateman et al. (2018). Importantly, an "[a]typical personality development can only be identified by considering the difficulties in negotiating developmentally appropriate, normative tasks that have relevance for the particular disorder of interest" (Fonagy & Bateman, 2007, p. 84). Due to social deviations from normative relational expectations (social norms) and low epistemic trust, such relational patterns tend to play out frequently in therapy with BPD patients (Morken et al., 2019). Therefore, the therapeutic relationship and focusing on repairing alliance ruptures seem central for positive outcomes for BPD patients (Paper III explored this further). In the course of treatment, the BPD patient will typically test whether the therapist can be trusted, for example, whether the therapist is willing to challenge them if they try to avoid painful content in what they describe. Therefore, becoming an attachment figure for patients implies tolerating the multiple relational tests while keeping a steadfast focus on the tasks and goals of treatment. However, this may induce difficult (counter)transferences and involve projective identification, for example, in terms of feeling a wish to abandon the patient or feeling invaded.

Being aware of such patterns may prepare the therapist for such emotional challenges in therapy. A somewhat parallel finding to that of Cologon et al. (2017) is that the

> higher the percentage of patients in a therapy group who had a history of relatively mature relationships, the better the outcome for all patients in the group, regardless of the form of therapy or the individual patient's quality of object relations score. (Piper et al., 2007, p. 116)

This also seems to signal that a sufficient ability to mentalize (RF) allows others (e.g., group members and therapists) to create a relational environment encouraging new emotional and interpersonal learning to happen. Therefore, it seems crucial for therapists working with BPD patients that they do some kind of self-developmental work to maintain their own RF when presented with poor mentalizing and prepare to tolerate the strong (negative) transferences typically involved in such treatments. A well-functioning (MBT) team is also necessary for such a process and for the carefrontational style allowing for successful battles of the comfort zone. Further, the merger of Zen Buddhism and cognitive therapy (CT) has proven useful for BPD treatments (Ellison, 2020). The integrating of such realms in therapy will be elaborated below.

The fourth major finding was that therapeutic alliance was associated with adherence and quality. The findings in Paper II suggest that quality/competence in reality (or at least in practical terms) is a measure of the specific alliance fostered by a focus on mentalizing (a specific alliance for MBT). As will be argued later, this could mean that future measurement of MBT and other psychotherapies should really be (or is) a specific measure of the working alliance promoted by the treatment approach. When rating CT, the mean correlation between adherence and competence has shown to be .96 (Barber et al., 2003, p. 214). This may signal that there is no significant difference between adherence and competence in CT. Further, it has been demonstrated that many raters struggle to distinguish between the two concepts (Perepletchikova, 2007, 2009). However, in the MBT-I-ACS no such confusion was reported (Karterud et al., 2013), and it seems that in MBT, perhaps because it is a relationally oriented psychotherapy, it becomes quite apparent to the observer/rater whether the interventions are successful in targeting the aim of the treatment in a way tailored to the patient. For example, as the manual states, the quality of the therapy reveals itself either as an attuned presence or as an absence in cases when, for example, the therapist is mistuned to the RF of the patient or when the therapist is not sufficiently challenging (Karterud et al., 2020).

Any clinical intervention is inextricably bound to the relational context in which it is applied (Strupp, 1986). Therefore, slavish adherence to treatment protocols appears to result in deterioration of the therapeutic relationship (e.g., Henry et al., 1993). Such inflexibility has been related to a negative outcome (e.g., Castonguay et al., 1996; Høglend et al., 2006) in that the therapist may try to fit the patient into a model instead of adjusting the model to the patient (Roth & Fonagy, 2006). Consequently, Owen and Hilsenroth (2014) emphasize the importance of therapist flexibility in relation to therapy outcomes. Contemplating the four sessions in Paper II, some may argue that the patients' contribution to the observed competence of the therapist is substantial. When Waltz et al. (1993) rigorously defined adherence and competence, they realized that the context of therapy characteristics of the client and what was happening in therapy were important. "When clients like their therapist and improve substantially, it is easier for therapists to look competent" (p. 624). Therefore, adherence and competence appear, in part, to be a function of the characteristics of the patient (Boswell et al., 2013; Imel et al., 2011). In fact, moderate adherence may imply therapist flexibility or responsiveness (Stiles et al., 1998), and it could therefore be viewed as a limitation that Paper II chose to investigate four sessions rated 2/2 and 7/7. However, as will be elaborated later, competence/quality may be seen as a (partial) measure of the working alliance (therapeutic alliance), and the two highly rated sessions in Paper II should not be seen as displaying rigid adherence but rather as examples of two therapists tailoring their method to the unique relationship. "In principle, each individual patient needs a unique form of treatment, adapted to her individual problems, needs, and style. This demands a high level of therapeutic flexibility, which most competent therapists are, however, able to offer" (Jørgensen, 2004, p. 519).

There have been some attempts at investigating the relationship between alliance and the specific technique. Gaston et al. (1998) found that the interaction between alliance and technique differentially predicted outcome between therapies. In short-term psychotherapy, 15% of the variance in outcome (measured as interpersonal problems) was explained by the interaction between alliance and exploratory techniques (not significant). In long-term psychotherapy, both supportive and exploratory therapist techniques interacted significantly with alliance to predict outcome. Further analyses indicated that supportive interventions were more helpful for patients with low levels of alliance, while exploratory interventions were more effective for patients with high levels. Further research in such directions may provide

more clarity on how good execution of MBT and other psychodynamic therapies includes but transcends the manual guidelines, which may in turn help improve future manual(s). Investigating 646 patients (9.5% PDs), Falkenström et al. (2013a) reported their "findings indicate that when the alliance is worse than usual for a given patient, symptoms are likely to get worse to the next session" (p. 326). As will be argued later, the finding in Paper II that the strategic competence and working alliance indicates the quality of the therapy and the importance for a constant "battle" in such challenging therapies may indicate a need for close monitoring of the alliance, session to session. This topic will be further addressed when debating how to teach, monitor, and manualize MBT. However, as the working alliance and strategic competence (not defined by the manual) seemed to explain the rated differences in MBT, one may wish to investigate the impact of the different aspects of the alliance in MBT therapies with different outcomes, which was the topic of Paper III.

5.5 Paper III

A recent study of 15,000 people receiving a range of psychological treatments across 184 services in England and Wales found that 5.2% of the patients had lasting negative effects from treatment. Patients who were unsure what type of therapy they received reported more negative effects (Crawford et al., 2016). This is perhaps not surprising, but still such findings underscore how important it is that patients "need to accept and engage in the therapeutic process not simply be engaged with the therapist but actively working toward a goal in a coherent way" (Wampold & Imel, 2015, p. 259), that is, focus on tasks and goals in therapy. In congruence with Paper II and indicating the importance of goal-directed work in MBT, Paper III highlights the importance of the working alliance (especially tasks and goals) in this therapy. The following four major findings will be discussed in depth:

1) Positive temporal development of therapeutic alliance during therapy characterized good outcomes.
2) Tasks and goals are particularly important.
3) Initial alliance ratings were high.
4) Comorbid paranoid PD is difficult but possible.

Our first major result was that MBT treatments with good outcome displayed a positive temporal development of therapeutic alliance. The importance of the alliance is particularly pronounced when treating patients with personality pathology (Benjamin & Karpiak, 2001;

Clarkin & Levy, 2004; De Bolle et al., 2010; Muran et al., 2009; Piper & Joyce, 2001). In fact, as mentioned earlier, the standardized effect of WAI-S on subsequent change was six times greater than in the group without personality problems (Falkenström et al., 2013a, p. 325). A stronger therapeutic alliance has been shown to predict greater improvement in BPD (Barnicot et al., 2012; Barnicot et al., 2011). For example, Marziali et al. (1999) tested how the therapeutic alliance in an RCT of interpersonal group psychotherapy and individual dynamic psychotherapy for BPD contributed to the outcome. The authors reported that alliance ratings were related to outcome in both individual and group therapy. Linehan (1993) suggested that a good relationship that is high in rapport is essential to treating clients with BPD because these individuals may be unable to fully utilize any other form of reinforcement to change behavior. Within the field of MBT, Fonagy and Bateman (2006) have introduced attachment as a core mechanism of change and state that concepts "such as the therapeutic alliance speak directly to the importance of activating the attachment system" (p. 411). Recently, the importance of alliance in MBT has been described in a qualitative study of MBT patients with comorbid substance abuse. Morken et al. (2019) state that

> [According] to our findings, good therapists know when to keep distance and when to come close, they are explicit about the content of own mind, they address the elephant in the room, and they tolerate strong affect. They put focus explicitly on the relationship between themselves and their patients. These findings resonate well with existing knowledge on therapist factors where the ability to form strong alliances and facilitative interpersonal skills is found to be essential. (Morken et al., 2019, pp. 10–11).

According to Bordin (1994), as therapy progresses the strength of the working alliance would build and ebb in the normal course of events, and the repair of these stresses in the alliance offers potent therapeutic possibilities and makes a direct contribution to clients' change. This is in line with the theory that negotiating the alliance and repairing alliance ruptures may be especially important in BPD treatment (e.g., Morken et al., 2019). Linehan has also argued that alliance problems are frequent and that their resolution can lead to the client's acquiring skills that can be used in interpersonal difficulties outside the sessions. Linehan's (1993) "techniques of acceptance" involve the therapist's ability to see reasonableness in the client's dysfunctional behaviors, accept the client's hostile affect, and recognize his or her own mistakes. Such an intervention style seems to foster (epistemic) trust, which may be particularly important when working with attachment disturbances. For instance, in working

with trauma victims, Hembree et al. (2003) have noted that trust is an absolutely essential element of the therapeutic relationship in prolonged exposure therapy because of the difficult and distressing nature of the process. However, despite attachment pathology being at core of BPD, we have found no quantitative studies investigating the relationship between alliance and outcome in MBT.

Could it simply be that "good" patients form good alliances and are destined to get better? Paper II indicated that experienced change would in turn lead to increased epistemic trust and improved working alliance. Therefore, observable symptom change may foster epistemic trust in the treatment and thus the patient's adherence to therapy (focus on tasks and goals). Several studies have demonstrated that the level of early alliance can be the product of previous symptomatic changes (e.g., Barber et al., 2000; DeRubeis & Feeley, 1990). DeRubeis et al. (2005) have presented an "alliance-as-outcome" hypothesis whereby the alliance is partially or even wholly an effect of previous symptom reduction rather than the cause of symptom reduction (e.g., Crits-Christoph et al., 2006b). "Thus, alliance is at least in part a pseudo-outcome, and the alliance–outcome correlation represents the correlation between two outcome measures" (Baldwin et al., 2007, p. 847). As we have seen arguments that a strong therapeutic alliance may be an appropriate therapeutic outcome for BPD patients (Barber et al., 2010), this could be particularly true for these patients. However, if

> alliance is a consequence of initial symptom change, even partially, then those patients
> with higher alliances should have better outcomes. In contrast, we observed that within
> therapists those patients with relatively high alliance ratings did not have better
> outcomes than those patients with relatively low alliance ratings. (Baldwin et al., 2007,
> p. 848)

Importantly, Tasca and Lampard (2012) have proposed a reciprocal influence model for the relationship between alliance and symptom change in which alliance and symptoms affect each other throughout treatment, and "our results support this model" (Falkenström et al., 2013a, p. 326). The transference "is a universal phenomenon of the human mind, it decides the success of all medical influence, and in fact dominates the whole of each person's relations to his human environment" (Freud, 1961, p. 42). Importantly, without the reference to any unconscious components, Bordin's (1994) working alliance captures the core curative therapy process conceived in Freud's classical concept of transference (Horvath, 2018, p. 504). One concept identified as the practical common denominator across alliance measures, is the "clients enthusiastic participation" (Hatcher et al., 1995) in treatment. Hence, the

findings in Paper II and III would be aligned with the model by Tasca and Lampard (2012), and it seems reasonable to assume that the patients' degree of enthusiastic participation would typically be associated with perceived improvement.

Let us approach alliance as "a way of looking at the relationship through the lens of effective goal-directed work" (Hatcher, 2010, p. 25). If this is indeed the case, one could argue that to establish agreement about the tasks and goals of therapy, the alliance is dependent on the delivery of a particular treatment (Wampold, 2019). However, in many manualized treatments, "the alliance has typically been treated as a factor that facilitates the use of and adherence to specific techniques not as a change mechanism itself" (Castonguay et al., 2010, p. 154). In addition to an alliance itself being a potential therapeutic outcome with BPD patients (e.g., a focus on alliance and alliance ruptures is in itself curative), we know little about the importance of the three subscales of alliance in psychotherapy in general and particularly in long term-treatments (Stiles & Goldsmith, 2010). There has been little research concerning alliance in (evidence-based) treatment for BPD, and due to "the paucity of research in this area, [Levy et al., 2010] had no reason to predict that different aspects of the alliance bond, agreement on tasks, and agreement on goals would be differentially affected" (p. 415). As Paper II indicated that the bond part of the alliance was an asset the highly rated therapists utilized to focus on tasks and goals, we became curious as to whether tasks and goals would characterize good MBT outcome better than bonds. Interestingly, in this respect Falkenström et al. (2015) state that tasks and goals may be one factor (not two), something which is in line with our findings in Papers II and III in that the development of Bonds seems somewhat different than tasks and goals. However, they conclude that because "meaningful differences between these scales are sometimes found in substantive research (...) it may be premature to conclude that the task and goal factors should be combined into one" (p. 591).

The second major finding in Paper III was that an increase in tasks and goals is particularly characteristic for treatments with good outcome in MBT. This is what we would expect to find from such an above-discussed interplay between symptom reduction and alliance, as tasks and goals are the parts of the alliance most connected to actual observable change (for the patient). In line with such an argument, Brotman (2004) suggested that therapists' "encouraging active involvement in their patients will improve patient adherence" (p. 35): Although encouraging active involvement is a general alliance-enhancing technique, here the focus is quite specific; patients should facilitate therapist adherence to concrete techniques,

and therapists should do whatever they can to encourage patients to participate in this way. Consequently, actual perceived/experienced improvement may help patients feel competent (efficacy), increase their trust in both the therapist and the method, and understand and collaborate on the tasks and goals of therapy, thus enhancing the probability of success and building the alliance further. In turn, a strong alliance will increase the likelihood that patients will agree with their therapists on the tasks and goals of therapy, which will affect outcomes (Baldwin et al., 2007; Wampold et al., 2007a). Therefore, when change occurs and an alliance is adequate, it makes sense that the alliance effect is especially strong when working with (B)PD patients (e.g., it can offset a synergy effect between increased epistemic trust, actual change, and increased agreement on tasks and goals).

The third major finding in Paper III was that the average initial alliance levels were above 4, which is considered satisfactory. Obviously, patients' attachment style and social competencies may affect their ability to foster a strong alliance with their therapist (Fonagy & Bateman, 2006; Mallinckrodt, 2000). Therefore, it was surprising that initial ratings were in the satisfactory range. However, we know that good alliance is created very early in treatment. In fact, it has been reported to be high before some patients even meet their therapists (Iacoviello et al., 2007). Further, given that all patients are required to participate in 12 psychoeducation sessions upon enrolment in the MBT program, one possible explanation for the initial high alliance reported in Paper III "is that it may be easier to follow the treatment in the first phase of MBT, in the period providing more structured psychoeducation" (Kvarstein et al., 2020, pp. 8–9). Importantly, the initial working alliance ratings—goals, bonds, and tasks—were not associated with outcome

The fourth major result in Paper III was that comorbid paranoid PD was more frequent in the subgroup with poor outcomes and associated with poorer alliance development in this subgroup. However, there were patients with paranoid PD in the good outcome group as well, and here, comorbid paranoid PD was not associated with poorer alliance development in the good outcome subgroup. This is an important clinical finding, as it imprints the therapist to have and display hope and work towards a common therapeutic goal. Somewhat similarly, Baldwin et al. (2007) argued that the "therapist attributions of resistance or maladaptive attachment styles as an explanation of a poor alliance, according to our findings, would be irrelevant with regard to outcomes, although these explanations may be grist for therapeutic work" (p. 851). Interestingly, in terms of attachment styles, or at least mentalizing abilities

(measured by Movie for the Assessment of Social Cognition; MASC), one recent study (Kvarstein et al., 2020) indicated that patients who tended to hypermentalize felt increasingly more able to bond with their therapist and find agreement on the aims and tasks of therapy. One may wonder whether a hypermentalizing style resembles an ambivalent (instead of a more dismissive/disorganized) attachment style. If so, that could provide explanatory power to the negative treatment effects often reported for avoidant PD (Kvarstein, 2013). However, one should be careful when associating alliance with personality categories (e.g., applying categorical variables to constructs deemed dimensionally distributed). Pedersen (2008) found that

> the number of personality traits defining these disorders were only slightly associated to the perception of the treatment milieu (…) it was the number of PD criteria, not the presence or absence of PD diagnosis that contributed to the variations in the perceived treatment milieu. (p. 72–73)

5.6 Measuring mentalization-based treatment

Taken together, what do these three jigsaw puzzles (the three papers in this study) tell us about measuring MBT, and how do they fit into the larger puzzle of psychotherapy research? We have seen that a positive alliance development characterized treatments with good outcome (end-GAF ≥ 60) in MBT (Paper III). Recall that Kvarstein et al. (2015) found a dropout rate of 5% in the first 6 months, while the typical dropout rate in similar studies has been reported to be much higher (Karterud et al., 2020). Alliance is one obvious factor for keeping patients in therapy. Specific "techniques have evolved over the years to deal with alliance difficulties" (Hatcher, 2010, p. 18), and while the best predictor of outcome is the alliance (Wampold & Imel, 2015), the best promotor of alliance is likely "good technique, technique that engages the patient in work that feels meaningful and goal directed" (Hatcher, 2010, p. 25). The indications that Tasks and Goals seemed more important for quality/outcomes than Bonds in Paper II and III resonates with such reasoning and also supports the idea that "alliance is not the same as the relationship (Hatcher, 2010, p. 25). Therefore, following our Ariadne's thread, I suggest that embedded alliance in MBT is focused on the goals and tasks of therapy (Paper II and Paper III) in a way that challenges the patient's comfort zone and consists of a strategy of treatment that is tailored to the best knowledge of the disorder, the unique individual, the relationship, one's own preferences, and the treatment method. Embedded alliance would capture and measure the fit between this tailored treatment and the patient, which would also include patient motivation, belief in the

method (both patient and therapist), the placebo effect, and personal chemistry. In MBT, an evidence-based treatment for BPD, epistemic trust seems to be an essential building block in the alliance (Paper II), specifically in terms of being the relational capital that seemed to allow for challenging the patient's maladaptive patterns of thinking, relating, and acting.

As our preluded conclusion (at the beginning of the thesis) alluded to, a specific focus on mentalizing (MBT) does not seem reasonable for most clients, and there is a need for a detailed recipe for MBT. This indicates that treatment integrity is also a crucial part of understanding the dialectic interplay of the mechanisms of change in psychotherapy (CF/EST), and Paper I indicated that both treatment components in the conjoint MBT program now have reliable integrity measures. However, competent (and actually also low-rated – for diametrical reasons) MBT contains interventions not defined in the manual (Paper II) and that seem difficult to transmit, largely because of the idiosyncratic nature of psychotherapy. However, investigating how the alliance is fostered in MBT is as close as we can come in terms of identifying, defining, and explaining such overall strategies. This idea should most likely be relevant for the larger field of psychotherapy research because there exists no CF treatment. Therefore, all therapies contain specific ingredients that need to be brewed as part of a larger "magic" potion, that is, a specific treatment method.

In Paper II, perhaps the most striking difference between highly rated and poorly rated MBT was the degree to which the curious, open, and not-knowing position was a vehicle for questioning, challenging, investigating, and identifying different perspectives in the patient's experiences. Consequentially, the more the relationship allows for contrasting views, confrontations, tolerating feelings, humour, self-disclosure, and curiosity, the more the patient seemed to improve their mentalizing capacity. Such features and activities seemed to reflect strategy and alliance, an alliance that needs to be particularly strong in MBT with BPD patients, providing a strong impact when present, as indicated by Paper III and Falkenström et al. (2013a). Simply stated, one could say that the "strong alliance" (patient and therapist agreeing that the goal is to improve mentalizing and be free from BPD) in MBT can be broken down, as light through a prism, into the 17 defining items in MBT-I and the 19 items in MBT-G. These ingredients can be seen as different ways to obtain an intact and efficient alliance in MBT (e.g., challenging, exploring in a not-knowing way, validating, acknowledging, displaying genuine interest, exploring the therapy relationship, stopping pretend mode, stopping psychic equivalence, being open and transparent about one's own

mind). "The unique aspect of MBT lies less in each individual item per se, than in the overall 'package' of item design and context" (Karterud & Bateman, 2010, p. 61). Therefore, as these 17 and 19 elements comprise a totality, addressing different cornerstones for change and preventing possible escape routes for the patient (e.g., "pretend mode"; pretending to be normal), it is important that the whole scale is applied (in a way tailored for the patient). This means that one can achieve a high adherence score in individual MBT without employing all the items in one session, but many missed opportunities (if indicated) would signify poor MBT (Karterud et al., 2020; Karterud et al., 2013).

Staying on course in the treatment (e.g., focusing on tasks and goals) may be particularly important with BPD patients. Levy et al. (2005) found that low RF in BPD patients predicted higher levels of impulsivity on the Continuous Performance Task and deficits in concept formation on the Wisconsin Card Sorting Test. Such impulsivity and vacillations "in mental states also lead to wavering on commitment to goals and in-session collaboration and demands from the patient for help followed by evasive maneuvers" (Levy et al., 2010, p. 415). Therefore, although some kind of intuition (e.g., non-linear process) may be a crucial denominator differentiating therapists' effectiveness in fostering a strong alliance, one should not advocate random interventions (towards unclear goals). For example, "decision making is a two-step process: First take in all the relevant information, then decide" (Dalio, 2017, p. 188). BPD patients inhabit a foggy and unclear inner reality, and it is necessary for efficient therapists to navigate MBT sessions with a normative map (Karterud et al., 2020). Kernberg (2016) has called for a unifying theory of PDs. A Temperament-Attachment-Mentalization-Based (TAM) Theory of Personality and Its Disorders has been proposed (Karterud & Kongerslev, 2019b) as one promising theory. Such topics will re-emerge later in the discussion. However, integrating a map is one side of the coin. Addressing corresponding issues in oneself is another factor likely to influence effectiveness. In Paper II, being trapped by one's own transferences seemed to be the major mechanism present in the low-rated MBT sessions, and working through one's own issues around being blamed, attacked, idealized, devalued, hated, ignored, or relationally tested in other ways could be of value for therapists working with BPD patients. According to Sehgal (2019), "Rumi emphasizes, 'Yesterday I was clever, so I wanted to change the world. Today I am wise, so I am changing myself'" (p. 39). This is in line with the atomistic approach of the current thesis. Hegel (2010) denied that thought could refer to unique individuals; rather, it is exclusively concerned with universals. Addressing the universality of human experience, Eliot wrote that the "business of the poet is

not to find new emotions, but to use the ordinary ones" (Kermode & Kermode, 1975, p. 43). Further, Buddhist philosophy (Brown, 2006) and even modern leadership theory seem to indicate that if one dives deep inside the self/particular, universal principles may appear.

> Look, if this really is your 'one-big-thing,' if you've really dug deep enough, if you've really gotten personal, everybody already knows. I know. Others know. So, there's this sort of illusion out there that you are sharing something so private, that nobody knows. Trust me, they know! (Kegan & Lahey, 2009, p. 81)

Consequently, self-exploration (in various arenas) may be an important source of achieving an inner map for how to skillfully navigate MBT sessions, for example, handling alliance ruptures. As will reappear in the further discussion below, learning different evidence-based treatments for BPD may also be advanced as a way to increase one's repertoire of helpful strategies. One central factor in all these treatments may be a psychopedagogic stance, and pedagogic interventions are suggested in the MBT manual (Karterud et al., 2020). Therefore, as an indirect pedagogy currently seems the most prevalent intervention in MBT (Folmo et al., 2021b), an elaboration concerning epistemic trust and pedagogics also seems relevant for the current thesis. This will reappear later, largely as a theoretical expedition (an empirical investigation is outside the scope of this thesis), because "there exist certain 'timeless truths,' consisting of common observations of how people change. These observations date back to early philosophers and are reflected in great works of literature" (Goldfried (1980, p. 996).

The "proliferation of treatment models and techniques results in a tremendously diverse set of healing practices that, at first glance, have little in common" (Imel & Wampold, 2008, p. 249). Therefore, the CF "approach to understanding the effects of psychotherapy holds that the unique theoretical content of an intervention is not an important guide to the mechanisms responsible for client change" (Imel & Wampold, 2008, p. 249). However, moderator analyses in a study by Høglend et al. (2006) indicated that specific treatments worked through different active ingredients for different patients, "implying that we can no longer expect simple causal explanations of treatment intervention and specificity of effect" (Bateman, 2007, p. 3). Consequently, the field of "psychotherapy is notoriously difficult to study" (Imel & Wampold, 2008, p. 249), and certain "treatment formats and settings may be helpful for some patients and yet damaging to others" (Bateman, 2007, p. 3). However, the CF approach is concerned with determining the core ingredients inherent in all successful psychotherapies so that a more parsimonious understanding of therapy can be developed

(Norcross & Goldfried, 2005). As we have seen that the field of MBT should embrace the CFs in order to develop further, what can the CF approach assimilate from investigating MBT? It seems like we complete a circle when claiming that good MBT-I creates a strong alliance, which is why it seems to work. However, an alliance is not an alliance, and each technique seems to produce different ones (Bordin, 1979, 1983, 1994; Falkenström & Larsson, 2017, p. 167). As will be discussed below, it seems that the CF approach may learn the four major themes from investigating MBT: 1) how to create an alliance in the presence of low epistemic trust and attachment pathology, 2) the importance of handling alliance ruptures, 3) the importance of being impeccable in one's own method, and 4) the importance of avoiding acting out own countertransferences.

BPD patients put pressure on the treatment system and demand therapists and teams who can co-operate (Bateman et al., 2012), build alliances (e.g., Paper III), and avoid acting out (negatively). "Individuals with borderline personality disorder are often unfairly discriminated against within the broad range of mental professionals because they are seen as troublemakers" (Lingiardi et al., 2005, p. 45). Traditionally, very high dropout rates (43%–67%) have been reported in PD samples of patients, mainly borderline, referred to psychotherapy (Lingiardi et al., 2005, p. 46). However, in a study of 33 BPD patients, Gunderson et al. (1997) found that therapist ratings of the alliance (after 6 weeks) were predictive of subsequent dropping out. Papers II and III reinforce the importance of alliance in MBT, and the carefrontational style advocated in Paper II is in line with the importance of focusing on tasks and goals indicated by Paper III. "Because people with personality disorders often try to test the limits of the therapist or professional when in treatment, proper and well-defined boundaries of the relationship with the client need to be carefully explained at the onset of therapy" (Lingiardi et al., 2005, p. 45). Further, Paper II indicated that low-rated MBT was hallmarked by therapists being overwhelmed by countertransferences. Therefore, therapists "need to be especially aware of their own feelings toward the patient, when the client may display behavior that is deemed inappropriate" (Lingiardi et al., 2005, p. 45). In terms of advocating certain interventions in evidence-based treatments for BPD, timing may be more important than the applied technique. Some studies report that transference interpretations were more helpful for patients with a lifelong history of less mature object relations (Høglend et al., 2006), while inverse relationships (the relationship differed as a function of quality of object relations) have also been found between the frequency of transference interpretations and both patient-rated therapeutic alliance and favorable outcome

(Ogrodniczuk et al., 1999). Importantly, even though different "psychological treatments, although apparently quite diverse, share much in common" (Imel & Wampold, 2008, p. 249), this does not imply that the "specific effects of psychotherapies may be substantially stronger than is widely believed" (DeRubeis et al., 2005, p. 174). Investigating outcome differences among therapy techniques' "'purity' provided significant correlations with outcomes (mean r=.44), both across therapists and within each therapist's case load" (Luborsky et al., 1986, p. 602). This may be seen as a result of therapists providing a pure potion and conveying a strong belief in their own treatment method, which is assumed to be a cornerstone for therapeutic change (Falkenström et al., 2013b, p. 10).

> The popularity of a therapy school is often a function of variables having nothing to do with the efficacy of its associated procedures. Among other things, it depends on the charisma, energy level, and longevity of the leader; the number of students trained and where they have been placed; and the spirit of the times. (Goldfried, 1980, p. 996)

However, having an overarching theory (school), method, and strategy may indeed foster such an allegiance, increased epistemic trust (e.g., via the three Communication Systems), and also a placebo effect.

Therefore, as the MBT manual (Karterud & Bateman, 2010; Karterud et al., 2020) suggests, it is probably the aspect of MBT in its totality that is the (most) curative element (i.e., the entire MBT program and the full spectrum of interventions). Further, according to such a view, one could argue that in being faithful to the manual and its mentalizing stance, the therapist displays an overall strategy (increases epistemic trust) that allows the patient to discover for themself (learning how to learn), stop non-mentalizing (pretend mode and psychic equivalence), focus on feelings, challenge erroneous views, and use their own countertransference in a transparent way to maintain the focus in therapy. In line with such an argument, and as most would probably have guessed, Paper II found that in highly rated MBT there were many MBT interventions (high ratio), and interventions are from the whole range of the scale (balanced). The lesson from low-rated MBT is that an overactive care system or the urge to maintain a "normal bond" with reciprocal positive feelings is a noteworthy factor in making therapists disadhere to the manual. Having lost focus on the 17 items and mentalizing, the therapists in those sessions were effectively outplayed by the patients. In low-rated MBT, the interventions come randomly without an overall strategy, which is both a result of and a reason why the therapy drifts into either supportive accompaniment or resigns to teleology, psychic equivalence, and pretend mode (teleology is arguably a special version

of psychic equivalence). However, there are some problems in this narrative, as both Paper I and Karterud et al. (2013) inform us that "psychic equivalence" and "pretend mode" have uncertain reliability. Expert raters have a hard time agreeing on these concepts, just as therapists evidently do. These concepts also had low reliability in a study on the training of Danish raters (Simonsen et al., 2019) in the MBT-I-ACS. As both these concepts are part of the core MBT theory, this is somewhat unfortunate because it could make the MBT potion even less specific (POT). Such speculations converge with the critique of MBT being too abstract (e.g., Sharp et al. (2020). However, they may also be somewhat premature, as the reliability was acceptable for psychic equivalence. Nonetheless, Paper I concluded that "In this case, it is unclear whether the group therapists delivered interventions for Items 15 and 16 that were poor and/or unclear or if the concepts of pretend mode and psychic equivalence were somewhat unclear for both therapists and raters" (p. 346). In terms of specificity, Paper I also showed that some PDG sessions achieved high MBT ratings. Therefore, there seems to be a strong (expected from the CF theory) but not complete overlap between MBT and other treatments, which seems to imply that MBT is a particular and quite distinguishable therapeutic "dialect" or language. However, in line with the argumentation above, it is also possible to suggest that MBT is first and foremost a specialized way to create a working alliance (focused on mentalizing) with a particular type of patient previously considered difficult to treat. The present thesis indicates the importance of merging adherence and competence/quality with alliance-measures tailored for the specific therapy (Hatcher, 2010), in this case MBT. Such a working model may allow MBT to maintain its language but also open the door for extensive collaboration with other treatment approaches. Importantly, the training of MBT therapists, supervision, manuals, fidelity measures, and clinical practice should be colored by an understanding of the importance of what is termed "embedded alliance". Importantly, as we will return to below, the current study highlighted the significance of epistemic trust in our interpretation of how the therapeutic alliance was nurtured. MBT should embrace the CFs, but it also seems that the broader psychotherapy community has much to learn from investigations of MBT, particularly in terms of epistemic trust and (natural) pedagogy. The phrase "two steps forward, one step back" may remind us of how this POT may synthesize some of the very essence that we are all looking at through different lenses. Despite attachment issues in the target population, there is little research on the relationship between alliance and outcome in the field of PDs. Investigating alliance development in manualized treatments, especially for patients displaying relational pathology, could potentially bring some more clarity on how different approaches bring about different

alliances. The current thesis has presented four major themes the CF approach may learn from the field of MBT. To the degree these topics reinforce current assumptions, this may reflect how the field of psychotherapy research can best grow

> [t]o the extent that clinicians of varying orientations are able to arrive at a common set of strategies, it is likely that what emerges will consist of robust phenomena, as they have managed to survive the distortions imposed by the therapists' varying theoretical biases. (Goldfried (1980, p. 996)

5.7 Through the looking glass

The mean alliance level for therapists has been reported to be a stronger predictor of outcome compared to differences in alliance between patients (Baldwin et al., 2007; Crits-Christoph et al., 2009; Falkenström et al., 2013a; Zuroff et al., 2010). Contemplating this conundrum, Lemma et al. (2011) repeats what Fonagy reasoned in 2010 (p. 36–37 [replacing "I" with "we"]):

> There is the legend of the therapeutic alliance, still frequently taught in (k)night schools. … More recent research that contrasted the outcome of patients with a number of therapists found that indeed differences between the effectiveness of therapists could be predicted by the strength of alliance they were likely to form with their patients (Baldwin et al., 2007) but differences in outcome between patients with the same therapist were unrelated to therapeutic alliance. If therapeutic alliance was the mechanism of change, then we would expect to do better with patients with whom we form a good alliance than those with whom my alliance is relatively poor. This turns out spectacularly not to be the case. (p. 17).

Ignoring the fact that "the lack of a within-therapist correlation can be understood from a rater bias perspective" (Baldwin et al., 2007, p. 850), one of the things we find most spectacular here is that, except in the study by Falkenström et al. (2013a), many of these studies seemingly do not include time (e.g., change trajectories or repeated measures of the alliance). Much of the research on alliance outcome is the product of the alliance being measured in Session 4 (e.g., Baldwin et al., 2007, p. 850; Wampold & Imel, 2015). To me, in general it appears somewhat problematic that the core mechanism of change in psychotherapy research is typically based on clustering "complex human relationships onto one evaluative dimension, called 'alliance strength'" (Stiles & Goldsmith, 2010, p. 45) and then often measuring it only once or twice over the entire course of the treatment before finally making inferences about

its causal impact (e.g., given that it is often research based on correlations). For example, in the study by Ulvenes et al. (2012b), patients rated the therapeutic alliance at the close of Sessions 4 and 20 using the Helping Alliance Questionnaire (HAQ; Luborsky et al., 1983), while HAQ scores were measured by Svartberg et al. (2005) in Sessions 4 and 20. Further, when investigating the alliance, "the total correlations—which in essence are a crude average of between- and within correlations—may provide misleading results" (Baldwin et al., 2007, p. 851). Discounting the implied imprecision of the involved instruments, it seems that our field is pervaded by various puzzling reports, such as that specific techniques account for less than 1% of the variance; researcher allegiance accounts for about 70% of the variance; extreme ranges in published therapists' effects; and statements like "differences among therapists in patient-rated alliance reduced the therapist effects by 97%" (Baldwin et al., 2007, p. 849). However, pyramids are not built by casting magic spells on random collections of stones (e.g., paraphrasing Poincaré, 1905), and the thesis began by asking why researchers in the field of psychotherapy have reached a schism in their interpretations, despite observing the same facts. "Scientists see new and different things when looking with familiar instruments in places they have looked before" (Kuhn, 1996, p. 111). One example of this is the above-mentioned argument that outcome equivalence between different treatments in RCTs does not imply that the same mechanisms produce the outcomes (DeRubeis et al., 2005). Laska and Wampold (2014b) stress that adherence to treatment protocols is not related to outcome and is characteristic of the patient. It might be the case that Wampold's interest in showing that the EST approach is a degenerative research program could in fact influence his own viewpoint and that his emphasis on the CFs, such as alliance, is somewhat biased. That said, the author of this thesis might well be bound by the same spell, so the reader must pay close attention to the argumentation to counter the influence of any possible attempt or wish to conclude that our present study makes (some) sense. However, let us focus on the alliance again. Applying multilevel models seemingly contrary to Lemma et al.'s (2011) attempt to cast a magical spell on (the legend of) the therapeutic alliance, in investigating temporal change trajectories for the working alliance both Falkenström et al. (2013a) and Paper III in the present study demonstrated that patients with a positive temporal alliance development had better outcomes. Therefore, it seems that it is the change in the alliance (slope) and not the possible rater-biased intercept that is associated with outcome. Paper III had a mean number of measurements of 3.2 and the initial ratings were all high, so a generalization of the findings could be somewhat restricted. However, just like Hegel replaced Aristotle's three logic laws by introducing time into them (Hoeflin, 2004), it seems somewhat misleading to

take a snapshot and then imply that we do not do better with the patients we have a positive alliance development with; that is, the alliance is not a static feature but a dynamic one.

5.8 Naïve and sophisticated falsificationism

In various misinterpretations of Kuhn, "science becomes merely one of numerous different readings of the text of the world, with no more actual authority than poetry, astrology, or palmistry: all are equally legitimate interpretations of the blooming buzzing confusion of experience" (Wilber, 2000b, p. 110). Hopefully, introducing the four quadrants (Wilber, 2000b, 2001a) and measuring MBT with methods anchored in all three of Habermas' validity claims should be sufficient to escape such naïve criticism. A particularly pronounced version of this approach is taken by the literary scholar and critic Harold Bloom (2014), who ranks Freud among the 26 magisterial figures of the Western canon: "Freud called himself a scientist, but he will survive as a great essayist like Montaigne or Emerson, not as the founder of a therapy already discredited (or elevated) as another episode in the long history of shamanism" (p. 12). Perhaps somewhat in line with such an argument, Hoffart and Johnson (2017) claim that psychodynamic explanatory concepts represent definitions of symptoms rather than causal explanations of them. Consequently, these models cannot be examined with regard to mechanisms of change (Cohen et al., 2018). For instance, reviewing a study by Luborsky (1984) where a depressed woman apparently recognizes the pattern behind why she is depressed, Hoffart and Johnson (2017) suggest that "the correlation between increased self-understanding and a decrease of symptoms of depression is not between independent phenomena and must be considered pseudoempirical" (p. 1079).

> In this view (…) wishes and emotions are independent of their behavioral manifestations. Similar to assessing whether a piece of metal is real gold by checking its number of protons, one could think that a final validation of the CCRT [The Core Conflictual Relationship Theme invented by Luborsky in 1975] formulation would be a detection of an inner force—an experienced wish for affirmation—in connection with social anxiety. However, this view that wishes and emotions terms stand for inner experiences represents a logical impossibility, most profoundly shown by Wittgenstein (1953). (Hoffart & Johnson, 2017, p. 1079)

As there are no certainties in interpretations, one may misunderstand Hoffart and Johnson (2017) at this point (e.g., they could mean that psychodynamic therapy is best investigated in the upper-left quadrant), but let us here interpret them literally according to the theory of

Wittgenstein (1953) and thereby allow their argument to serve as an example of confusing different validity claims (Habermas, 1986); "exterior surfaces can be seen, but interior depth must be interpreted" (Wilber, 2000b, p. 184). Collapsing the interior domain (upper left) into the upper-right quadrant, our only valid interpretation of Shakespeare's *Hamlet* would be to give a precise account of the number of protons or other physical particles in the original manuscript (Wilber, 2000b, 2001a). Applying the theory of mentalization at this point, one may be inclined to suggest that even Harold Bloom would disagree with such disregard for his entirely interpretational professorship (upper-left quadrant). In terms of psychotherapy research, it may be of interest for our discussion regarding different validity claims that according to Hoffart and Johnson (2017), contrary to the change mechanisms inferred by psychodynamic pseudoscientists the "causal schemas of CT [cognitive therapy], on the other hand, provide micro-networks of causal relations between circumstances, activities, and experiences" (p. 1079). Based on Euclidian geometry, the authors explain why this is so, stating that

> If an object is geometric and has three sides, then it is a triangle. And if it is a triangle, then it is geometric and has three sides. (…) In definitions, linguistic entities are associated with conditions in natural reality. Causal relations, on the other hand, associate events in natural reality with other events in natural reality. (Hoffart & Johnson, 2017, p. 1078)

Despite the danger of this being "researcher allegiance in favor of CT" (Leichsenring et al., 2018, p. 19), this sounds promising for CT. However, in terms of mathematics and interpretation, Roger Penrose (Emeritus Rouse Ball Professor of Mathematics at the University of Oxford/Cambridge and Nobel Laureate in Physics) seems to indicate that when "we convince ourselves of the validity of Gödel's theorem we not only 'see' it, but by so doing we reveal the very non-algorithmic nature of the 'seeing process itself'" (Penrose, 1989, p. 541). Consequently, the idea that the human mind operates via algorithms, as a computer does, in a wholly machine-like, rule-driven manner would mean that we could never perceive the truth of this theorem because it is not derived algorithmically (which the proof proofs). Evidently, even the measurements of protons imply interpretation according to present physics (various interpretations of quantum theory, e.g., Penrose, 2006). Indeed, it seems that

> [i]nterpretation is not something added onto the Kosmos as an afterthought; it is the very opening of the interiors themselves. And since the depth of the Kosmos goes 'all

the way down,' then, as Heidegger famously put it, 'Interpretation goes all the way down.' (Wilber, 2000b, p. 184)

In line with the importance of different validity claims, Wittgenstein argued that philosophy is to provide a kind of therapy enabling us to correct fallacies of thought, or more precisely, that philosophical problems are misconceptions that are to be therapeutically dissolved. As will be explored later, this resonates with the importance of a pedagogic stance and epistemic trust in psychotherapy. However, more importantly in our current context, he also makes it clear that "there is not a single philosophical method, though there are indeed methods, different therapies, as it were" (Wittgenstein, 2009, §133d).

However, most importantly for the present attempt to measure MBT, a psychodynamic treatment, Hoffart and Johnson (2017) argue that psychodynamic therapy does not meet the second criterion proposed by Kazdin (2009) for psychotherapy research, which is a strong association between the proposed mechanism and therapeutic change. Therefore, without demonstrating a relationship between RF and outcome, MBT would have little or no scientific value. Hence, in light of statements like "it is as yet unclear whether specialist and nonspecialist treatment models are equally effective in bringing about change in RF in conjunction with symptomatic improvement" (Chiesa et al., 2020), we could contemplate whether a useful research question for the field of MBT would be if the theory of increased RF/mentalization can be falsified? Lakatos would indeed argue that this is the wrong question: "While naïve falsificationism stresses "the urgency of replacing a falsified hypothesis by a better one, sophisticated falsificationism stresses the urgency of replacing any hypothesis by a better one" (Lakatos et al., 1980, p. 37). Science can grow without any "refutations" leading the way. "Naive falsificationists suggest a linear growth of science, in the sense that theories are followed by powerful refutations, which eliminate them; these refutations in turn are followed by new theories" (Lakatos et al., 1980, p. 36). Thus, two ESTs could easily have the same fundamental mechanism of action in reality, despite contrary theory (i.e., the theory is wrong). Further, the fact that a theory of change mechanism might be unfitting does not by itself invalidate the worth of the therapy. "For the naïve falsificationist science grows through repeated experimental overthrows of theories; new rival theories proposed before such "overthrows" may speed up growth but are not absolutely necessary; constant proliferation of theories is optional but not mandatory" (Lakatos et al., 1980, p. 37). Importantly, however, "a bad interpretation of *Hamlet* is falsifiable, not by sensory data, but by further mental data, further interpretations – not monological data but

dialogical data – generated in a community of interpreters" (Wilber, 2000b, p. 221), for example, the epistemological basis of the G-study performed in Paper I.

5.9 Proliferation: Marriage of common factor and empirically supported treatment

The honesty principle of sophisticated falsificationism demands that one should try to look at things from different points of view to put forward new theories that anticipate novel facts and to reject theories that have been superseded by more powerful ones (Lakatos et al., 1980, p. 38). In line with this, some psychotherapy theory and research of late has focused on how the CF and the specific ingredients work together to produce the benefits of therapy (Hoffart et al., 2012; Owen & Hilsenroth, 2011; Pesale et al., 2012; Ulvenes et al., 2012b; Wampold & Budge, 2012). Much of this research shows that the CFs do their work differently in different therapies. Rather than argue about which perspective is more or less scientific, we must ask how can one integrate the two models of empirical inquiry in a way that the field can move forward? Laska et al. (2014) state that

> The ideal training program, in our view, should contain elements of both the EST and the CF perspectives. Psychotherapy trainees should be trained to provide ESTs as well as be trained in feedback systems, and how to form and repair strong working alliances, express empathy, and collaborate on treatment goals. We recommend going one step further, however, to provide "competency-based" certification. That is, to be certified, trainees would need to attain outcomes with various types of patients that meet a given standard (e.g., the benchmark for a particular disorder or type of patient). (p. 476)

The current thesis fits well with such a recommended path forward, and this will be elaborated in a later paragraph presenting suggestions for how to teach MBT, directions for future research, and implications for the Quality Lab for Psychotherapy Research. However, before we can talk about effect, we should investigate what the outcome, reflected in tasks and goals, in therapy may mean.

5.10   Outcome in psychotherapy

Both day hospital and outpatient treatment modalities have been recommended as cost effective for patients with Cluster B disorders in the Netherlands (Soeteman et al., 2008). Here in Norway, Kvarstein et al. (2015) reported very high effect sizes (i.e., 1.4–1.7 on symptom reduction and interpersonal and social functioning). As mentioned earlier, MBT has been demonstrated to have a superior effect in RCT, which is an important finding because

most treatments for BPD have traditionally failed (i.e., treatments predating evidence-based treatments for BPD; Karterud et al., 2020). However, other comparisons of specialized borderline PD treatments have also reported clinically significant improvements but failed to reveal the superiority of any specific approach (Clarkin et al., 2007; Jørgensen et al., 2013; McMain et al., 2009). As will be discussed later, both therapist and researcher allegiance may play a part here; for example, Elkin (1999) argued that different teams and their culture may have a larger effect than the prescribed treatments. Therefore, there is a risk of confounding treatment efficacy with therapist (or team) efficacy. Elkin (1999) puts significant effort onto creating competing teams with comparable characteristics, competence, training, and supervision. However, it should be noted that psychotherapy is not only about treatment effects and efficacy. Importantly, there has typically been little patient participation in defining outcome measures (Duncan et al., 2010). Therefore, the field should try to find ways to continually ensure that general well-being; global assessment of life (e.g., purpose in life); and having a partner, friends, work, and a social network are included in our evaluation of treatments.

Frank and Frank (1993) suggested that the "aim of psychotherapy is to help people feel and function better by encouraging appropriate modifications in their assumptive worlds, thereby transforming the meanings of experiences to more favorable ones" (p. 30). Next, despite the very welcome term *clinically significant change* and later advocated by Lambert and Ogles (2009), we should not forget that within the upper-left quadrant there is also an important aspect of therapeutic change that implies that change does not always imply feeling better. According to Bion (1977),

> All helpful endeavours have a foundation which is, like most foundations, unobserved – the belief that things can be improved. Even psychoanalysis is tainted with ideas of cure that imply a better state. I think it is 'better' to know the truth about one's self and the universe in which I exist. But I do not wish to imply that it is 'nicer,' or 'pleasanter.' Whether it is 'better' is a matter of opinion which each individual has to arrive at for himself: his opinion and only his. (p. ii)

Ignoring his post-Hume criticism of falsely assuming universally imprinted concepts and presuming to legislate what other people meant (e.g., displayed poor mentalizing and/or disrespect for the hermeneutic tradition) or should have meant, Kant (2007) wrote, that "[t]o coin new words is to arrogate to oneself legislative power in matters of language, which is

rarely successful" (p. 297). Nonetheless, Kiesler (1995) coined the term "the myth of patient uniformity" (p. 94) in referring to the belief, prevalent in psychotherapeutic theory and research "that all patients with a given diagnosis are fairly uniform on all the factors that are relevant for psychotherapeutic treatment, and that they thus require the same treatment" (Jørgensen, 2019, p. 57). In his argument for a differential epistemology, Johansen (2008) argues that the first "I" and the second "I" are not identical in Descartes' "indubitable" statement "I think therefore I am" ("je pense, donc je suis"; 1960), and it seems that psychotherapy researchers have clustered patients together with patients, often ignoring the fact that each patient is indeed a unique individual and is the final arbiter in any clinical setting (Habermas, 1986, p. 261). The thesis will return to this point and the need for patient participation in defining outcome measures when presenting possibilities for future research. However, let us now take a closer look at the very basis for evidence-based practice, the manual.

5.11   Spelling out the word spell: Are manuals helpful?

Attempting to approach the upper-right quadrant, providing a treatment manual is supposed to make it possible to standardize the independent variable by reducing therapist differences in RCTs and outcome studies (Crits-Christoph et al., 1991). Not surprisingly, this process of defining a treatment seems to have been more successful for highly skills-based therapies (Sharp et al., 2020) than for psychodynamic treatments, where the core of the therapy is being highly attuned to idiosyncratic variables that will determine how best to intervene. As Paper II demonstrated, highly rated MBT contains overarching strategies not defined by the manual. Therefore, given that the very essence of the working alliance is to tailor the treatment to the unique relationship (Bordin, 1979), how can we merge the need for treatment guides (pronouncing principles of the therapy model advocated) with the ability to implement the approach flexibly and skillfully? In other words, as each individual is unique and teaching others to mentalize has been criticized for being too abstract (Sharp et al., 2020), how can we teach it without providing "endless" examples? Let us here take one step back and assume that Wittgenstein (1975) was wrong and that there is a "second order spelling". Next, let us translate this philosophical idea in terms of a data program, so that the spelling of second order is rather the ability to investigate any set of program rules and predict what happens when the program is run. Can we, or a computer, investigate the rules of a computer program (no matter how small, simple, complicated, or large) and then predict what happens when it executes? Given a program with, let us say, only two naïve rules, would we be able to predict

what happens without simply running/executing it? Contrary to what most of us would probably expect, this was proven by Gödel (1931), Turing (1936), and Langton (1986) to be impossible, that is, even with artificial intelligence (AI) and/or quantum computers and no time limit. At this point, the observant reader will naturally wonder if this has any practical implications for understanding psychotherapeutic treatment manuals. Gödel, Turing, and Langton talk about logical systems and not human interactions. However, when we learn that predicting what happens from two simple rules is mathematically proven to be impossible, then we can perhaps climb up this Wittgensteinian (2001) ladder of understanding and then silently throw away the ladder. From a logical standpoint then, we know that it is absolutely impossible to predict what happens in, for instance, MBT, no matter how many principles or guidelines we cultivate and define. Consequently, even an infinite MBT manual would not be very helpful (not only because it would be impossible to read).

It has been suggested that the problem (e.g., "manualized therapy provoking varying degrees of allergic reaction" and raising "concerns about the mechanisation of the therapeutic process") is not inherent to manuals per se but rather relates to how they are used and that the presence of a manual can serve the function of some supportive superego. "In this respect having the manual in mind supports a certain discipline in clinical decision making because it forces us to think explicitly about why we may decide to do something other than what is advocated in the manual" (Lemma et al., 2011, p. 23). Evidently, in many instances not following the manual is indeed indicated. Castonguay et al. (1996) reported that

> a unique aspect of cognitive therapy (i.e., therapist's focus on the impact of distorted cognitions on depressive symptoms) correlated negatively with outcome at the end of treatment (…) Such increased focus, however, seems to worsen alliance strains, thereby interfering with therapeutic change. (p. 497)

Rigid adherence to manuals and prescribed techniques has been found to deteriorate the therapeutic relationship (Henry et al., 1986; Henry et al., 1993), for instance by neglecting more pressing problems in the relationship, such as an empathic failure (Burns & Auerbach, 1996). The same is true for too blindly focusing on incomplete or outdated case formulation (Beck et al., 2015; Persons, 1989). A study by Barber et al. (2006) indicated that for patients with low alliance, adherence to the treatment model was necessary for their improvement, whilst patients with high alliance improved independent of the therapist's adherence. Further, ruptures in the alliance have been found to be linked to rigid adherence (Castonguay et al.,

1996; Henry et al., 1993; Piper et al., 1991). Therapists typically tried to resolve ruptures by increasing their adherence to a theoretical model, such as challenging distorted cognitions in CT or providing transference interpretations in dynamic therapy. These studies found that high adherence in the context of an alliance rupture was linked to poor outcome and premature termination. "However, there is also evidence that training therapists in manualized approaches that emphasize the formation and maintenance of a strong alliance may improve some therapists' abilities to manage alliance ruptures successfully" (Eubanks-Carter et al., 2010, p. 80). Crits-Christoph et al.'s (2006b) initial excursion into alliance-fostering psychotherapy indicated that moderate to large improvements from pre- to post-training on alliance occurred (not statistically significant).

Recall the statement that "the treatment method is the only aspect in which psychotherapists can be trained" (Chambless et al., 2006, p. 199). Crits-Christoph and Connolly (1999) also argued that some consistent findings relating techniques to outcome indicate that nonspecific relationship factors do not by themselves account for the changes found over the course of short-term dynamic psychotherapy. However,

> [i]f one assumes that these multiple factors are in constant interaction and interdependence in ways that sometimes enhance and sometimes hinder change, one is indeed forced to recognize that the complexity of the psychotherapy process goes beyond a debate between 'specific versus nonspecific' or 'techniques versus relationship.' (Castonguay & Grosse, 2005, p. 201)

In line with such an argument, Ogles et al. (1999) concluded that "little evidence substantiates the benefit of technique-based training" (p. 215) in their review of existing empirical studies. It is also stressed that a "[k]nowledge of therapeutic strategies and techniques does not guarantee that a therapist will be competent" (Lemma et al., 2011, p. 23). We have also seen that therapists who, on average, formed stronger alliances with their patients showed statistically significant better outcomes than therapists who did not have as strong alliances (Baldwin et al., 2007, p. 849), which could be taken as an argument to study talented therapists (e.g., Lemma et al., 2011). Further, Boswell et al. (2013) observed that "Over half of the variance in adherence and competence was explained at the session level, suggesting that fidelity is contextually driven" (p. 451). Consequently, based on our findings in Paper II and Paper III, it is possible to argue that adherence and competence (quality) can be merged, measured, and perhaps better understood by the term embedded alliance, which also implies

that quality (competence) can be seen/understood/recognized and perhaps rated as the interplay/combination between adherence and the working alliance.

5.12 Embedded alliance: A consequence of the marriage

By 1994, more than 1,000 process-outcome findings involving the alliance had been reported (Orlinsky et al., 1994). Yet, "in comparison to the large number of studies examining the relation between alliance and outcome, there are relatively few examining the impact of the Alliance×Technique interaction on outcome" (Barber, 2009, p. 2). However, one study comparing alliance in ST and TFP (Spinhoven et al., 2007) indicated that factors specific to a particular approach influence the quality of the alliance. ST with its emphasis on the "necessary and sufficient conditions" as identified by the client-centered approach (Rogers, 2012) produced a better alliance according to the ratings of both therapists and patients. Hatcher and Barends (2006) noted that "alliance and technique occupy different conceptual levels and cannot be considered to be two different types of activity in therapy. Technique is an activity, alliance is a way to characterize activity" (p. 294). Based on our findings, it seems reasonable to advocate the view that the alliance is embedded within each therapeutic method and cannot exist without specific technique(s). It seems understandable/logical that the definitions of the working alliance have been numerous and often confusing. As mentioned, Crits-Christoph et al. (2014) stated

> [o]ur view is that the "CF perspective" should be subject to the same sorts of empirical investigations as any other "perspective" on behavior change. Thus, rather than an "alternative," the "CF perspective" has the potential to be an "EST" as well, assuming that the "CF perspective" can be described in a way that provides guidance to therapists about how to conduct therapy. (p. 491)

Interestingly, in some forms of psychodynamic treatment (Strupp, 1984) the management and therapeutic use of the alliance is viewed as the core technique of intervention. In hindsight, Henry and Strupp (1994) seem to have argued, somewhat prematurely, that their construction of a treatment manual centered on the effective management of the alliance implied that the schism between the specific and non-specific factors was fixed. The future will tell whether the current dissertation falls in the same trap of underestimating the temporal force in the universe.

Kant coined quite a few concepts himself, but the current thesis will reason that the term "embedded alliance(s)" (Bordin, 1979; Hatcher, 2010) will do to describe what is arguably the core of what the present study implies in terms of tailoring the specific treatment approach to the unique patient (Jørgensen, 2004, p. 519). In fact, in many cases moderate adherence may imply therapist flexibility or responsiveness (Stiles et al., 1998). However, high quality/competence may arguably be seen as the degree to which the alliance is embedded within the treatment. Investigating MBT-G, Esposito et al. (2020) found that "it is not so fundamental that the clinician is completely consistent with the treatment manual. Rather, and more importantly, clinicians should use adherent interventions competently" (p. 9). Therefore, they "claim that clinicians' skills in adapting treatment to the clients and context might be more effective than close adherence to a treatment manual per se" (ibid.). Observing the correlation between adherence and quality in Paper I and the fact that it may be difficult to deliver MBT interventions of "high quality" (Esposito et al., 2020, p. 1) without adhering to MBT (as we do not rate the quality of interventions that are not part of the MBT model), one may wonder if what Esposito and colleagues try to communicate is that MBT quality is really a measure of embedded alliance. At least, this is what one may propose based on our current study (Paper II). Therefore, as different treatments demand different alliances (Bordin, 1979), this may indeed necessitate new ways of assessing "magic potions". "Measurement of specific alliance features can be guided by examining the 'embedded alliances' that exist in particular therapies" (Hatcher, 2010, p. 15). Such measurements of embedded alliance remain largely unexplored but could shed some light on reports that "patients who facilitated therapists' adherence to concrete techniques demonstrated significantly more improvements in the following session". In practical terms, this could mean "patients who were able, interested and/or willing to provide specific examples of events or cognitions" (Brotman, 2004, p. 33), that is, signaling that an adequate working alliance and adherence are interdependent. In terms of such an embedded alliance, "[i]t appears likely that for particular treatment approaches some components of the collaborative work will have greater influence on the outcome or may even be critical to success" (Hatcher, 2010, p. 15). Therefore, applying an EST, "[t]he therapist can learn how to use or take advantage of the common factors more deliberately and can use more specific technical interventions selected for the problems of the individual patient" (Jørgensen, 2004, p. 536). Paper II (and perhaps also Paper III) supports that it is not the form or manual in itself but each therapist's capacity to tailor those guidelines to the unique patient. The guiding principle is recognition that psychotherapy is implemented one person at a time based on each unique individual's

perceptions of the progress and fit of the therapy and therapist (Duncan et al., 2010). This also means that good therapy will include but transcend the manual, as indicated by Paper II.

5.13 Your potion is only as strong as you believe it is

All human civilizations have practiced some form of healing with designated healers, shamans, elaborate explanations for illness, and healing rituals (Wampold & Imel, 2015, p. 52). Therefore, humans have developed placebo treatments to heal through social means that "contribute to a large extent to the success or failure of patient treatment, sometimes even more so than the drugs available" (Enck & Zipfel, 2019, p. 9). For instance, Khan et al. (2008) reported that "only 1 out of 5 of those that improved with placebo relapsed when continued on placebo for three months or more" (p. 795). As both CFs and specific factors contribute to the placebo effect in psychotherapy, Enck and Zipfel (2019) conclude that "[i]t is now time for psychotherapists to accept them in their daily practice" (Enck & Zipfel, 2019, p. 9). A wide array of CF models have been proposed to explain the benefits of psychotherapy, many of which include different versions of the placebo effect as a central curative ingredient. According to (Lambert, 1992), four therapeutic factors influence the outcome of psychotherapy: a) client/extracurricular factors; b) relationship factors; c) placebo, hope, and expectancy factors; and d) model technique factors. Frank and Frank (1993) discussed six elements that are common to the rituals and procedures used by all psychotherapists, the second of these being that the therapist maintains the patient's expectation of being helped by linking hope for improvement to the treatment ritual. Kiesler (1986) "proposed that in the early stage of successful therapy, it is inevitable for a therapist to be 'hooked,' i.e., to provide a complementary response to the client's maladaptive interpersonal behavior" (Dolan et al., 1993, p. 408). The third component is that there exists a rationale, conceptual scheme, or myth that provides a plausible explanation for the patient's symptoms. According to Frank and Frank (1993), the particular rationale needs to be accepted by the client and the therapist but need not be "true." The rationale can be a myth in the sense that the basis of the therapy need not be "scientifically" proven. "Using the medical definition of placebo, the effects of psychotherapy are ipso facto placebo effects, and psychotherapy is ipso facto a placebo" (Kirsch et al., 2016, p. 121).

> However, it is critical that the rationale for the treatment be consistent with the world view, assumptive base, and/or attitudes and values of the client or, alternatively, that the therapist assists the client so that he or she is in accord with the rationale. Simply

stated, the client must believe in the treatment or be lead to believe in it. (Wampold & Imel, 2015, p. 48)

In modern western cultures, psychotherapy is acknowledged (and often formally approved by the authorities) as a practice aimed at relieving emotional suffering. This cultural recognition brings the necessary social legitimacy to the therapeutic rituals and to the communicated rationale or myth that offers the patient an understanding of her psychological problems. Simultaneously, this social legitimacy will support the patient's faith in the therapist and the therapy. (Jørgensen, 2004, p. 535)

The final component is a treatment ritual or procedure that requires the active participation of both client and therapist and that is consistent with the rationale previously accepted by the client. Contrary to the idea that RCT gives an objective answer regarding the efficacy of a treatment, one might argue that there are no pure observations but rather observations couched in a theory-laden vocabulary. Feyerabend (1970) argues that the observation language is part of the theoretical language rather than something self-contained and independent. Further, even when a therapy has been shown to be responsible for change in general (RCT), other factors than therapy may cause apparent reported changes. Therefore, "[e]ach of the existing psychotherapeutic rationales and theories is in part a social construction that is embedded in a specific cultural logic and form of society. Each therapeutic rationale allows the patient to order and explain her experiences" (Jørgensen, 2004, p. 534). As we have seen, "[f]or the rationale to be effective, both therapist and patient must accept it" (Wampold & Imel, 2015, p. 48), and it is obviously also crucial that the therapist adheres to the treatment and instills hope. This is in line with the finding that the highly rated therapists in Paper II did not abandon the therapeutic project and expressed explicit confidence in the method. In fact, the degree to which the therapist believes in the treatment is such an important factor that it may directly impact the outcome of psychotherapy. "Therapist allegiance remains a crucial unstudied factor in psychotherapy research. Strength of belief in a therapy may affect the therapist's comfort and authenticity in conducting treatment, the therapy's plausibility for the patient, and thus the strength of the therapeutic alliance" (Falkenström et al., 2013b, p. 10).

Therefore, congruent with arguments from the CF approach, Paper II indicated that expert therapists (high rating) provided a rationale for their therapy and had a clear belief that this

was an effective treatment approach. Despite this example being taken from a setting where the temperature was very high and the patient resisted using all means, this finding is in line with the guideline that "[a]ll evidence-based psychotherapies provide a coherent framework that enables the patient to examine the issues that are deemed to be central to him or her, according to a particular theoretical approach, in a safe and low arousal context" (Fonagy et al., 2019, p. 5). However, perhaps offering ammunition to those who would argue that Freudian psychoanalysis offers but another mystical explanation "in the long history of shamanism" (Bloom, 2014, p. 12), Allison and Fonagy (2016) argue that the experience of knowing and having the truth about oneself known in the context of therapy is not an end in itself but is important because the trust engendered by this experience (epistemic trust or trust in new knowledge) opens one up to learning about one's social world. Psychotherapeutic models differ in detail, but they generally work—directly or indirectly—to develop strategies to handle how one thinks and feels with regard to oneself and to restructure thinking about interpersonal relationships (Fonagy et al., 2019). Importantly, "[i]n order to safely depend on others to learn about reality, we need to be able to identify those who are reliable sources of information" (Fonagy et al., 2019, p. 4). We therefore need to investigate the role of pedagogic stance and epistemic trust in psychotherapy in general, but as we have seen above this activity may be crucial in MBT (Paper II). This is further indicated by the importance of goals and tasks (Paper III) and the reported importance of psychoeducation in MBT (Ditlefsen, 2020).

5.14 Pedagogic stance: Some clinical implications of epistemic trust

Goldfried (1980) suggested two principles common to all therapies—providing the client new and corrective experiences and offering the client direct feedback. This is in line with the finding in Paper II that the highly rated therapists challenged the patients' comfort zones, which seemed to foster epistemic trust. All evidence-based treatments for BPD contain strong elements of pedagogy and strategies to achieve epistemic trust. For instance, ST utilizes cognitive, experiential, and behavioral interventions and focuses heavily on the provision of a strong, quasi-parental relationship between patient and therapist (Ellison, 2020). However, as we have seen in Paper II, in highly rated MBT this is exemplified by balancing a curious "not-knowing stance" with a communication of central psychological building blocks (or challenges):

> To be a teacher is not to say: This is the way it is, nor is it to assign lessons and the
> like. No, to be a teacher is truly to be the learner. Instruction begins with this, that you,

the teacher, learn from the learner, place yourself in what he has understood and how he has understood it, if you yourself have not understood it previously, or that you, if you have understood it, then let him examine you, as it were, so that he can be sure that you know your lesson. (Kierkegaard, 1998, pp. 46–47)

Importantly, in conjoint therapies such as MBT an essential aspect is that other patients can communicate knowledge, demonstrate their own epistemic trust in the therapists, and "intervene" when therapists miss opportunities. Karterud (2015) has suggested that borderline patients will be more open to confrontations with their fellow patients than with an authoritative therapist, a tendency that may reflect difficulties with authority in general (e.g., low epistemic trust). To hopefully increase this effect, reading the case formulations to each other in MBT groups has been introduced as part of the treatment program in MBT (Karterud, 2018).

When entering therapy, the patient is at some level open to receiving help and gaining new knowledge about the world. The teaching aspect of the therapy is highlighted in the focus on epistemic trust. This term addresses the ability of some therapeutic "potions" to open the patient to a necessary trust in someone else's knowledge about the world. Thus, epistemic trust is an essential part of the "strong alliance" and is the element that defines the therapeutic relationship as something other than a normal, safe relationship. Paradoxically, however, as Plato teaches us (Hamilton et al., 1961), the most powerful and likely lasting way to teach is to allow for a not-knowing exploration (because the patient then learns how to learn and can trust his new knowledge at a deeper level). If a patient displays too much trust in the therapist, this would be a collapse in mentalizing, and the patient would need to integrate the tool of mentalizing by practicing together with the therapist.

The two low-rated MBT sessions in Paper II seemed to demonstrate that the countertransference of being useless, judged, not knowing enough, or not being liked, which resulted in attempts to establish a "normal bond" with reciprocal positive feelings was perhaps the most important factor making the therapists stray from the manual. The highly rated MBT therapists seemed to tolerate the patient's anger, irritation, not knowing, and stubbornness in a steadfast way, while the more poorly rated MBT sessions were characterized by therapists wanting to deliver solutions, trying to please, offering extra sessions, or avoiding difficult feelings and thoughts. We have learned from the MBT manual

(e.g., Karterud & Bateman, 2010) that the ability to withhold one's own opinions/knowledge/answers is essential for the patient's understanding and change. The therapist must have a plan for their questions (some Socratic method perhaps or whatever other pedagogic school one adheres to). This means that the therapist has in mind what the patient needs to understand but withholds answers as long as the patient is heading in the "right" direction (but not eternally withholding answers). Within the MBT model, the therapist might, however, give hints or tell the patient how they would have reasoned if in the patient's shoes (e.g., Item 15 "Use of countertransference").

In terms of epistemic trust, the difficulty in therapist training is twofold: 1) It is very difficult to imagine what it is like not to know something when you know it. However, the willingness to enter not-knowing with the patient is crucial because the patient typically needs to discover for themselves in order to fully trust their understanding and to change their maladaptive patterns. 2) It can be difficult to trust that someone actually is able to discover (learn) by themselves when they move slowly, that is, the therapist thinks their function is to mentalize for the patient. Perhaps the most important thing patients learn in MBT is to refine the tool with which they need to find answers by themselves (improved mentalizing and epistemic trust). This then clarifies that MBT therapists are teachers of a sort who teach people how to learn. One illuminating finding by Rønnestad and Skovholt (2003), who interviewed expert therapists, is therapists' ability and willingness to learn from their patients. The patients in MBT are ideally engaged in a mentalizing discourse where beliefs, feelings, and interpersonal transactions are challenged in order to bring about changes in perspective, while solutions and answers play subordinate roles (Karterud and Bateman, 2010). There is a significant difference between openly exploring what goes on in another's mind and evaluating statements about the world that can be tested. Therefore, the therapist needs to challenge unwarranted beliefs (Item 3). Psychic equivalence is a state in which the patient's thoughts become too real. This is also a state that needs to be challenged; however, it will often be necessary for the therapist to validate the patient's emotions first (Item 13 "Validation of emotional reactions") and then attempt to untangle the patient's view (Item 9 "Psychic equivalence").

As we have seen, Item 16 ("Monitoring own understanding and correcting misunderstanding") is the most frequent intervention in MBT-I and therefore deserves attention. It also works in symphony with most other items and is often closely related to Item

2 ("Exploration, curiosity, and a not-knowing stance"). Still, checking one's own understanding is not quite the same as pure curiosity, so the distinction is necessary. Changing one's worldview is considered a challenging process, not to mention amending one's personality, as most of us tend to believe what we believe. In trying to define insight in therapy, Hobbes stated that "The best definition I have been able to come up with is this: Insight is manifested when a client makes a statement about himself that agrees with the therapist's notions of what is the matter with him" (1962, p. 742). In such instances, Item 7 becomes particularly relevant.

Every teacher knows that pupils must be rewarded and encouraged on the way to increased knowledge. In Paper II, we saw several examples of the excellent use of Item 7 ("Acknowledging positive mentalizing"). This intervention seems crucial in building a strong alliance (epistemic trust), as it directly points to the therapist's approval of the patient's endeavor to learn mentalizing. The essence here is that the therapist remembers what the patient does not know and steps out of their own understanding without forgetting what the patient is looking for. And when the patient finally achieves understanding, the therapist joins the (tea) party. Item 7 is not only linked to Item 2, but in some cases the acknowledgment could be directed towards the patient's willingness to mentalize or to tolerate feelings. Improved functioning is also an expression of mentalizing that should be acknowledged. We saw no concrete examples of this in our four sessions. This absence was "penalized" in the competence ratings of the low-rated session, but the absence of Item 7 was not seen as a mistake in the highly rated session. This may be due to a halo effect, incompleteness in the manual, or simply that the acknowledgement comes across even though it is not explicitly verbalized. One example might be the therapist smiling in session B as the patients says she is angry with her.

All evidence-based psychotherapies provide patients with a model of the mind and an understanding of their disorder and then treatment actions consistent with that explanation. However, any therapeutic model—that is, understanding the causes of the problem and their possible resolution—can be effective only insofar as it results in the patients' feeling of being mirrored in a way that leads to the feeling of being understood as an agent. In our view, this is one of the most powerful human experiences leading to the restoration of feelings of agency and selfhood (Fonagy et al., 2019). However mysterious (e.g., an expression of the universal will; Hegel, 2018, p. 234) our personalities may be, a Danish study has concluded that the

development of agency through the reconstruction of personal life stories may be a crucial mechanism in psychotherapy with BPD patients (Lind et al., 2019b). In line with this, another Dane has brilliantly written:

> If One Is Truly to Succeed in Leading a Person to a Specific Place, One Must First and Foremost Take Care to Find Him Where He Is and Begin There. This is the secret in the entire art of helping. Anyone who cannot do this is himself under a delusion if he thinks he is able to help someone else. In order truly to help someone else, I must understand more than he—but certainly first and foremost understand what he understands. (Kierkegaard, 1998, p. 45)

There are seemingly (at least) two important aspects here. First,

> If I do not do that, then my greater understanding does not help him at all. If I nevertheless want to assert my greater understanding, then it is because I am vain or proud, then basically instead of benefiting him I really want to be admired by him. (Kierkegaard, 1998, p. 45)

As "to help is a willingness for the time being to put up with being in the wrong and not understanding what the other understands" (ibid.), then not-knowing is essential in a treatment. Curiosity is an essential component of this process; we have termed this the not-knowing or inquisitive stance (Fonagy et al., 2020, p. 1). In terms of pedagogic stance, which has recently been suggested as an MBT intervention (Karterud et al., 2020), the overarching strategy will be in accordance with the patients' level of insight into their inner reality. For example, congruent with Kierkegaard (1998), Kohut (1984) states that the healing aspect of the analyst's interpretation is that the patient feels understood. To battle comfort zones in a tailored fashion, the therapist first and foremost needs to listen (e.g., inhabit the presented narrative).

> In the end, we would think of only one universal therapist stance: it seems to be important with all clients to be a good listener. Listening includes a focus on the client and communicating a sense of respect and interest in what the client has to say. Such therapist behavior can help to enhance the client's sense that he or she has been understood. (Dolan et al., 1993, p. 408)

The second crucial aspect is that the helper must indeed have more knowledge than the person they are trying to help (in that domain); that is, "I must understand more than he" (Kierkegaard, 1998, p. 45). The fact that appreciation and empathy are crucial in all helping endeavors—for example, "My love, if anything will help him to become another person"

(Kierkegaard, 1995, p. 172)—does not exclude the fact that true knowledge also has great importance and value; for example, "[t]he Sophist demonstrates that everything is true" (Kierkegaard, 2013, p. 205). "Truth—or, more precisely, an accurate understanding of reality—is the essential foundation for any good outcome" (Dalio, 2017, p. 135). Therefore, one should not recommend the frequent use of myths (adhering to the Sophists) in therapy but rather the best and most relevant (tailored) pieces of knowledge available. Wittgenstein summarized this well, stating "[t]o convince someone of the truth, it is not enough to state it, but rather one must find the path from error to truth" (Wittgenstein et al., 1993, p. 119). "If you can do it, if you can very accurately find the place where the other person is and begin there, then you can perhaps have the good fortune of leading him to the place where you are" (Kierkegaard, 1998, p. 46). Importantly, as the concept of pretend mode in the framework of MBT highlights, a mere mental understanding is not enough, as "[h]aving to exist with the help of the guidance of pure thinking is like having to travel in Denmark with a small map of Europe on which Denmark is no larger than a steel pen-point, indeed, even more impossible" (Kierkegaard, 1992, p. 275). Apparently, even wizards have no potion against emotional suffering, and at the affected end of the novel *Harry Potter and the Goblet of Fire*, Professor Albus Dumbledore provides us with a plausible psychoeducational model of how the mind works: "He will stay, Minerva, because he needs to understand. Understanding is the first step to acceptance, and only with acceptance can there be recovery. He needs to know who has put him through the ordeal he has suffered tonight, and why" (Rowling, 2000, p. 680). This kind of (grand)fatherly intervention is in line with what we found in the highly rated MBT sessions (Paper II), and the emotional content and sincerity of the therapist combined with the bond made the interventions potent. Importantly, "[i]f what the therapist offers in this respect is not felt to be true, the channel for knowledge transmission will remain closed, and the patient will be unable to learn from the experience of therapy (Allison & Fonagy, 2016, p. 298). Therefore, one potential pitfall in the focus on epistemic trust or natural pedagogy (Csibra & Gergely, 2006, 2009, 2011) is that it presumes that epistemic trust is a biological feature (upper-right quadrant; Karterud et al., 2020) and that as long as therapists succeed in activating the epistemic highway all is well. One may perhaps suggest that it is more important to investigate how and in which setting what content should be provided to the patient instead of focusing solely on such explanations and suggestions as ostensive cues (Sperber et al., 2010) that signal to patients the relevance to them of the information being conveyed (i.e., ostensive cues can be seen as signals designed to trigger epistemic trust; Csibra & Gergely, 2009). Examples of content that could be transmitted from therapist to

patient in an adaptive, timely, and pedagogic manner is such information that "[s]ecure dependence and autonomy are two sides of the same coin, rather than dichotomies […]. The more securely connected we are, the more separate and different we can be" Diana (Fosha et al., 2009, p. 263), or that we humans can change attachment patterns, e.g., "earned-secure attachment" (Dimaggio et al., 2007). The first one would fit well with a typical intervention addressing the second prototypical version of targeting impaired epistemic trust (Mistaking or confusing others' dependency, gratitude, or relational valence with reality). However, as such a pedagogic stance can color and be part of all other interventions, an empirical recognition of effective pedagogic strategies may necessitate the study of expert therapists. As discussed above and exemplified in Paper II, both Folmo et al. (2021b) and Karterud et al. (2020) reported a frequent use of Item 16. As 32% of the interventions in 327 MBT-I sessions were of this type, we get an empirical signal that there is something here in need of further investigation. One should be careful not to suggest an increased pretend mode in therapy; but based on the current study it seems that opening the door to a pedagogic stance in MBT is indicated. For those who believe in another myth, namely in worldly success (of which Ray Dalio, legendary founder of *Bridgewater Associates* is an example), for some patients in a suitable setting (etc.) such interventions could be carefully administering the content in principles like "Don't confuse what you wish were true with what is really true" or "Don't worry about looking good – worry instead about achieving your goals" (Dalio, 2017, p. 162).

As we have seen, the term *specific factors* is generally meant to refer to the core, theory-specified techniques or methods prescribed for a given treatment modality (Holtforth & Castonguay, 2005). While the MBT therapist tries to regulate feelings through the mentalizing discourse, when emotions reach the boiling point, the skills manual for DBT (Linehan, 2014) advises as follows: "Suppressing emotion increases suffering. Mindfulness of current emotions is the path to emotional freedom" (p. 403). However, if the emotional pain reaches dangerous/extreme levels, one should "[s]ay: 'Splash your face with cold water or put your face in a bowl of ice water or cold water on your eyes and upper face (this will reduce arousal for a brief time)'" (p. 402). Linehan reported that this surprisingly simple but effective technique will calm you down immediately. The theory here "might seem counterintuitive, but research has shown that immersing your face in very cold water while holding your breath causes your body to turn on the nervous system's relaxation response and slow your heart rate" (McKay et al., 2019, p. 110). However, when emotions are not at the boiling point, we could perhaps remind the patient that "you're looking for the best answer, not simply the best

answer that you can come up with yourself" (Dalio, 2017, p. 189) or to recognize the idea "that to gain the perspective that comes from seeing things through another's eyes, you must suspend judgment for a time—only by empathizing can you properly evaluate another point of view" (Dalio, 2017, p. 189). Importantly, "some skills might not be appropriate in every situation even if they do work. For example, it might not be practical – or safe – to use the diving technique while you are driving a car" (McKay et al., 2019, p. 132). In an integrative contribution to the understanding of mechanisms of change in BPD, Euler et al. (2019) reported that overall defense function (measured by the Defense Mechanism Rating Scale) in an RCT with 32 patients (16 of 31 outpatients received DBT skills training in addition to individual treatment as usual) improved significantly due to skills training in DBT. However, in line with Hoffart and Johnson (2017), caution in interpreting such studies is advised. Decreases in the defense mechanism targeted by many psychodynamic treatments may not add empirical support to DBT, not only because such concepts are suggested to be unmeasurable but also because DBT is "not explicitly designed to target defense mechanisms" (Euler et al., 2019, p. 1074). However, despite such warnings it seems that such findings point us towards a common ground for evidence-based treatments for PDs.

In terms of transmitting skills or other wisdom, there has been a recent surge in interest in the Eastern philosophies in Western mental health care. For example, Falkenström (2010) reported an increase in mindfulness was associated with an increase in well-being with the successful incorporation of Buddhist principles into DBT and the increased psychotherapeutic application of mindfulness (Bhatia et al., 2013). A recent pilot study by Schanche et al. (2020) concluded that "it may be beneficial to introduce personal mindfulness practice as a way of preparing novice therapists for their future profession" (p. 311). Such mergers of old philosophical and psychological teachings from Buddhism with Western psychology has been fruitful (e.g., Falkenström, 2003; Safran, 2003). This is not only important in terms of implementing knowledge but also in terms of the cross-cultural perspective whenever we try to investigate universal psychological principles. The Bhagavad Gita can serve as another example worthy of attention in relation to the current thesis. The Bhagavad Gita is an ancient collection of writings that can best be compared to the Bible or the Quran (except that the Hindus also include such works as the Relativity Theory by Einstein, The Holy Bible, and Shakespeare's writings among their holy scriptures). This work serves as the primary spiritual guide for the vast majority of Hindus worldwide (Bhatia et al., 2013). Numerous eminent Indian psychiatrists have recommended the use of the principles in the Bhagavad Gita for

psychotherapy and healing (Venkoba & Parvathi, 1974). In fact, it is considered the first "psychotherapeutic manual" by some, as researchers have identified most Western psychotherapeutic approaches (e.g., grief emancipation therapy, mindfulness, psychotherapy, psychodynamic psychotherapy, and supportive psychotherapy) embedded in this work produced around 2500 to 5000 BC (Bhatia et al., 2013). Throughout the work, the therapist (Lord Krishna) is a trusted friend, philosopher, and guide—one who certainly instills hope and utilizes ostensive cues as he says to his patient (Arjun) "Leave everything and trust me, I will rescue you from all the problems, do not worry" (Reddy, 2012, p. 102). A Western psychotherapist should most likely refrain from instilling so much hope and dependency (despite the patient [Arjun] being considered as having satisfactory premorbid personality with adequate coping skills; ibid.), but the point here is to see the cross-cultural aspects of communicating as a treatment method and identify epistemic trust (knowledge) (culturally embedded but still valid as recognized, tried, and true principles) within a trustful relationship (i.e., the working alliance).

> As is the case with any successful model of therapeutic intervention, which needs to be individualized for maximum benefit, the psychotherapeutic approach practiced in the *Bhagavad Gita* also will have its place in the repertoire of psychotherapeutic models and remains a useful tool in the hands of an experienced therapist when applied judiciously for some patients with specific problems of distress. (Reddy, 2012, p. 104)

At this point, it seems clear that

> [T]here is no such thing as a single true theory of psychotherapy and its active ingredients, nor is there one superior technique that can be applied to all forms of pathology, although specific techniques and curative factors may be particularly important in working with certain types of pathology. (Jørgensen, 2004, p. 534)

5.15   A common ground for evidence-based treatments for personality disorders

Wittgenstein (2001, § 5.6) famously stated that "The limits of my language mean the limits of my world", something Hoffart and Johnson (2017) also alluded to, and to the degree that the language or the logos we operate within is a large contributor to how we view our reality (Passer & Smith, 2004, p. 305), it seems reasonable to assume that this has profound impact on the field of psychotherapy. Observing the same patient, some perceive object relations, while others see schemas, negative thoughts, or as Tomkins (1992) put it, "the world we perceive is a dream we learn to have from a script we have not written" (p. 239). Different

therapies build their strategies around affect integration (Falkenström et al., 2014; Solbakken et al., 2011), cyclical psychodynamics (Wachtel et al., 2005), object relations theory (Fairbairn, 1954; Guntrip, 1973; Winnicott, 1956), maladaptive schemas (Young et al., 2006), prementalistic representations of internal states (Bateman & Fonagy, 2016), or maladaptive behavior patterns (McCullough et al., 2003). Other theories, such as the theory of primary emotional systems (Panksepp, 2011; Panksepp & Biven, 2012), have been operationalized in the Affective Neuroscience Personality Scales (ANPS) and have explained 20% and 19 % of the variance in borderline and avoidant criteria, respectively, in 546 patients with different degrees and qualities of personality pathology (Karterud et al., 2016). In short, the field of psychotherapy is inhabited by different and typically competing world views, paradigms, language, and treatments. However, many of these theories overlap in various ways. Patients with BPD display disturbances in self and other understanding, which is also evident when they narrate events from their own and significant others' lives (Lind et al., 2019b), which reminds us of object relations theory, scripts, cyclical dynamics, and other more or less famous models of the mind. Therefore, the recent striving for a sort of CF approach within the ESTs is liberating, that is, the idea that CFs support a dimensional conceptualization of PDs (Bateman et al., 2018, p. 44). Obviously, there should be a variety of traditions in order for different therapists and patients to find a "treatment myth" they can believe in and adhere to, but one may also suggest that as all these different theories and ESTs work in similar ways, a common basis of understanding would enhance future treatment programs for severe PDs. A recent comparison by Gunderson et al. (2018a) of different treatment strategies (and accompanying theories) for BPD calls for unifying theories of personality and PDs. Otto Kernberg's (2016) outline in his significant paper "What Is Personality" has been suggested as such a common theoretical ground (Karterud, 2017; Karterud & Kongerslev, 2019b). As the current thesis seeks to merge CFs with EST, it supports Laska et al. (2014) in their idea of a "competency-based" certification. However, one would not necessarily think that to be certified trainees would need to attain outcomes with various types of patients that meet a given standard (e.g., the benchmark for a particular disorder or type of patient) but rather to demonstrate competence in at least two different ESTs for a given disorder. This may sound like a harsh requirement, and the reader may wonder what the author may be thinking. Well, I am looking into the future, where *personalized medicine* will be a dominating trend. Some may even argue that it is even unavoidable (Porter & Teisberg, 2006). Therefore, I believe that such certified therapists should learn to create a working alliance by different methods. Psychotherapists should be curious. When one has read Plato, Dante, and Homer, then

perhaps one should read the Upanishads. And after one has memorized, or perhaps battled, Linehan's extensive DBT manuals, perhaps one should read some Bion. We should not lose sight of the details, however. To paraphrase the psychiatric patient and great mathematician John Nash (1956), the only person to be awarded both the Nobel Memorial Prize and the Abel Prize, simple representations of specific ideas are better but often more difficult to produce than less specific and more elaborated ones (every Riemannian manifold can be isometrically embedded into an Euclidean space). As we know from any pseudoscience, everything can be explained given an infinite number of "dimensions". "Open-mindedness doesn't mean going along with what you don't believe in; it means considering the reasoning of others instead of stubbornly and illogically holding on to your own point of view" (Dalio, 2017, p. 189). A reasonable perspective of the CFs approach to treatments of PD may prove helpful for understanding how the embedded alliance works, or in other words expand our vision of how the alliance interacts with different specific factors designed for the same diagnosis. What implications does embedded alliance have for deliberate practice in MBT?

5.16   How to teach, monitor, and manualize mentalization-based treatment

Mentalizing is difficult to master (and teach) because when attached to a patient's emotional state, it is hard to maintain one's own capacity to skillfully mind minds (mentalizing). Therefore, the effective MBT therapist will master the balance between being in tune with and being too engaged with the patient, the content/narrative, themselves.

> Instead, through focusing on and engaging with mental states underpinning actions, the therapist helps the client to generate multiple perspectives to free the client from one or more non-mentalizing modes, including the teleological mode (physical action is seen as the only way to modify someone else's mental state), pretend mode (the mental world is experienced as decoupled from external reality), and psychic equivalence (a mind-state where the distinction between the contents of the mind and the external world is unclear). (Sharp et al., 2020, p. 2)

MBTs must be able to engender a robust strategy to implicitly incorporate mental states spontaneously and effectively in a variety of social actions. In line with this tenet, we grapple with the question of how to teach people to mentalize without explicitly teaching them and without relying heavily on expert supervisors who can translate dense psychodynamic-based theory into practice, thus learning to "mentalize from the inside out" (Sharp et al., 2020). In Paper II, we discovered that therapists pursued goals according to their strategic competence.

Manuals have perhaps succeeded in describing the key ingredients of the potion (Sharp et al., 2020) but have not been able to instruct therapists in how to find the right balance between the ingredients or how to serve it, that is, how to follow overarching strategy. MBT is theorized to work by "restoring a balance between the different polarities of mentalizing (automatic versus controlled, self versus other, internal versus external, cognitive versus affective), by the therapist maintaining a 'mentalizing stance'" (Sharp et al., 2020, p. 2); however, few examples are provided.

Although we understand that "[k]nowledge of therapeutic strategies and techniques does not guarantee that a therapist will be competent" (Lemma et al., 2011, p. 23), manuals are needed in order to deliver highly specialized treatments such as MBT. However, perfectionist ideas of how therapy should be performed or how change happens in therapy can contribute to self-criticism and doubt for (novice) therapists (Rønnestad & Skovholt, 2003). In Paper II, the highly rated therapists demonstrated high self-confidence, and if one strategy failed they persistently tried another, aiming for the same goal(s). Beutler (1999) advised focusing "on learning to apply strategies rather than either broad theories or specific techniques" that are "designed to capitalize on therapist flexibility and clinical judgment in treatment" (p. 403–404). In terms of manuals, strategic competence, and navigating towards a goal in the presence of another mind, I find that chess is actually a good metaphor for understanding the principle of no perfect move or strategy. The author is aware that not everyone is as dedicated to the game of chess as he is, but as everyone knows, this ancient game exists somewhere between science, sport, and art. Here, I will argue that it is first and foremost an exhaustive exercise in mentalizing on many levels simultaneously. Be that as it may, what we can learn from moving around on a chessboard is that there is no right move, no one strategy to rule them all. With supercomputers, one can sometimes get this impression (and it may be the actual case in the endgame in chess), but in practical play against a flesh-and-blood opponent, there are endless strategies towards the same goal. Chess instructions for certain openings contain typical plans and obvious traps to avoid along the way in attempting to acquire some advantage, be it temporal, material, positional, and/or tactical. And as elsewhere, some of us are much better than others at mentalizing, in this case mostly due to chess being a language best learned when very young. However, we can still learn some strategies and general principles, such as displayed by Alexander Alekhine (the fourth World Chess Champion). In a tournament game played in Dresden in 1936, he noticed his opponent had set a clever trap for him. If Alekhine captured a certain pawn, his crafty competitor would reply with a move that

won a piece. After the game, Alekhine wrote "White falls into the trap," white referring to himself, "and thereby proves that it is the fastest way to win!" The creative reader may well imagine ways to generalize such a principle to the art of MBT. Informed by this metaphor and inspired by DBT, one could suggest that the MBT manual should perhaps contain different strategies for how to meet typical reactions to different items, to challenging unwarranted beliefs, and to therapists' frequent countertransferences to these. For example, if a therapist challenges such a belief and meets resistance of some sort the manual should give some guidelines as how to adequately react by bringing awareness to the (partially) subconscious tendency to accommodate the patient, hence the term "battles of the comfort zone" (Paper II). The same would most likely be true for pedagogic interventions ("Pedagogic stance"; Karterud et al., 2020). Therapy, even MBT, is not about not knowing. Based on Paper II, I propose that this Item 2 should be called "Curious stance" instead, as it seems that many therapists misunderstand what a "not-knowing stance" means.

> Contrary to our expectations, we found the typical chaotic borderline group processes that MBT-G was designed to avoid. An in-depth examination of the data, employing qualitative Thematic Analysis, revealed that the therapists failed to establish themselves as authoritative leaders of the group and misconstrued the 'not-knowing stance.' (Inderhaug & Karterud, 2015, p. 150)

Bordin (1979) offered no predictions concerning temporal fluctuations of the alliance, and most researchers have prioritized the association between early alliance and outcome (Stiles & Goldsmith, 2010). For short-term therapies, investigations of the shape of alliance processes over the course of short-term psychotherapy, including high average, U-shaped, and linearly increasing alliance processes, have been associated with clinical improvement (Stiles & Goldsmith, 2010). However, when it comes to the treatment of BPD, the alliance between each session may be crucial. After the low-rated sessions presented in Paper II, it seems reasonable to assume that it would be essential for the course of therapy that aspects of the therapeutic relationship, including the working alliance, be addressed in the subsequent session(s). As of yet, there is no session-to-session measure of the alliance in MBT. It has been suggested that the "session-to-session effect of the alliance on symptom level points to the importance of continually monitoring the alliance throughout treatment, perhaps especially if that patient has personality problems" (Falkenström et al., 2013a, p. 326). Monitoring the alliance closely may provide several benefits for clinical practice, research, and supervision, and it has also been found that systematic client feedback improves

effectiveness and legitimizes psychotherapy services to third-party payers (Duncan et al., 2010). Such a measure could be tailored to the specific treatment. As will be argued in the section concerning future research, based on the embedded alliance it seems that fidelity measures and manuals focused on the specific working alliance intended by each specific EST would also be advantageous. For instance, Bender (2005) concluded that transference interpretations should be used sparingly with the more disturbed PDs (e.g., those with borderline and narcissistic features) during the early phase of treatment, a cautionary outlook previously recommended by (Kohut, 1984). This is supported by the findings of Piper et al. (1991) for a mixed diagnostic group, namely that a higher proportion of transference interpretations is associated with a poorer alliance and that for patients with a poor quality of object relations, the close correlation of transference interpretations with dynamic formulation is predictive of a poor therapeutic alliance. According to Bender, supportive, empathic interventions are initially better (Messer & Wolitzky, 2010, p. 110). Importantly, in MBT there is a hierarchy of interventions (Karterud et al., 2020), and the timing of more complex interventions, such as the use of countertransference, may necessitate careful preparation (e.g., sufficient working alliance). This is in line with previously mentioned studies indicating the negative effects of such interventions "Inverse relationships were found between the frequency of transference interpretations and both patient-rated therapeutic alliance and favorable outcome. The relationships differed as a function of the patient personality characteristic known as quality of object relations" (Ogrodniczuk et al., 1999, p. 571).

In MBT, just like in music or chess, innumerable scenarios can emerge from the limited set of rules (interventions) or, as we have seen, even from two simple rules in a computer program (Langton, 1986). There are many pitfalls in MBT. One is mentalizing on behalf of the patient. "The risk is the illusion we are well able to create for ourselves that our mentalizing the patient is sufficient" (Allison & Fonagy, 2016, p. 286). Many of these challenges cannot be learned except through practice. Therefore, perhaps the most important lesson to learn from chess or music or anything else is that, in general, practice is the key to success. Disturbingly, our profession is one of few we are able to think of where one is encouraged to apply one's theoretical skills in a real-world setting before having practiced similar tasks in real life.

> A novice musician would not take courses in music theory and then go directly to performing on stage. Yet, the predominant format of psychotherapy training requires the trainee to go directly from coursework to seeing real patients. No matter how

talented the supervisor, the supervision of real patients only minimally approximates

the requirements of "deliberate practice." (Binder & Henry, 2010, p. 210)

As has recently been advocated, deliberate practice should be implemented in the training of therapists (Rousmaniere, 2016). As we will see shortly, this also has implications for the Quality Lab for Psychotherapy in Oslo.

5.17 Implications for The Quality Lab for Psychotherapy

User manuals can be a bit tricky, and the alleged allergy to such written instructions may or may not be correlated to why I know few people actually read the user manual after buying new equipment (although it is strongly recommended). We saw in the introduction that practice tells researchers where knowledge is most needed, so that science has its feet firmly grounded in everyday clinical care. This is crucial for the Quality Lab for Psychotherapy. Therefore, we should research what kind of feedback is most useful for therapists and (preferably) the actual connection between the feedback and what patients experience. Do the MBT ratings agree with what the patient experienced in the session? Would the patient agree with the rating? Such questions are extremely hard to investigate due to various ethical concerns, but perhaps a creative researcher can find a way to have a sneak peek anyway. For example, Inderhaug (2013) applied the methodology of member checking (Lincoln & Guba, 1985) to ensure therapists were able to appreciate his interpretation of the data (i.e., that they largely agreed with the researcher).

The current thesis suggests that the lab should integrate more measures, such as the Achievement of Therapeutic Objectives Scale (ATOS; McCullough et al., 2003) and PQS. Several researchers (Crits-Christoph et al., 2011; Ulvenes et al., 2012a) have shown that several psychotherapy process measures can be dependably rated simultaneously. Therefore, including the ATOS and the PQS as part of the rating could give important insight into the CFs (both are pan-theoretical instruments to measure important therapeutic processes). This may be an important step towards a more specific alliance-based (embedded alliance) measure for MBT. The lab (and other researchers) should also rate and investigate other therapy methods, such as TFP and affect consciousness therapy.

5.18 Future research and development: The greater puzzle

Despite attachment issues in the target population, there is little research on the relationship between alliance and outcome in PDs. Further investigations of alliance development in

manualized treatments, especially for patients displaying relational pathology (e.g., BPD patients), could potentially bring some more clarity on how different approaches bring about different alliances. The importance of future investigations of pedagogic interventions has also been strongly indicated (Folmo et al., 2021b; Karterud et al., 2020) by the current dissertation. Further, as it was the structuring element of the MBT-G-AQS (Items 1, 2, and 3) that was the major difference from PDG (Paper I), the impact of this structuring of sessions (treatment) should be investigated. One possibility would be to investigate PDG and MBT-G sessions qualitatively. As the current dissertation has focused on alliance in individual therapy, it seems like a natural next step to investigate the alliance in (MBT) groups. Alliance to the entire group has been indicated as central (Lindgren et al., 2008) and may provide scaffolding for future research. The conjoint aspect of MBT should also be investigated, as little is known about the connection between these core components.

In terms of teaching, operationalizing, and conceptualizing MBT, it is interesting to note that Bateman and Fonagy (2009) argue that MBT demands minimal training and supervision because it incorporates generic therapeutic principles and has a commonsense view of the mind. Contrary to this assumption, it has been noted that "treating borderline patients in groups seems to be extraordinary challenging, and that MBT-G seems to require extensive training and supervision" (Inderhaug, 2013, p. 40), which resonates with the previously discussed critique by Hutsebaut et al. (2012) and Sharp et al. (2020). Consequently, investigations of how therapists absorb the MBT model and what is needed to implement MBT as a team and as an individual or group therapist are much needed.

More research into both patient and therapist (or perhaps team) factors is strongly needed. Castonguay (1993) voiced concern that focusing on therapist actions ignored other common aspects of psychotherapy. One of many alternative models is that the patient is the most important CF and that it is clients' self-healing capacities that make therapy work (e.g., Bohart, 2000). This may be a bit presumptuous when dealing with BPD patients, but it may indeed necessitate a stronger need to investigate what works for whom; it seems plausible that individuals with different pre-treatment levels of mentalizing capacity may differ in their ability to engage in psychotherapy (Katznelson, 2014). However, one could also interview patients who have been treated with TFP, ST, DBT, and MBT and investigate how they assume change occurred or did not occur. IPA would be a possible choice of method for such a study. In this respect, when investigating these evidence-based treatments further, I fully

support the CF approach, taking the view that more RCTs investigating minor differences is simply the wrong focus (e.g., Messer & Wampold, 2002).

Another important missing piece in the jigsaw is whether we can produce reliable measures for the overall structure of the MBT program and whether such a rating scale would correlate with treatment fidelity and/or outcome. After hearing one piece of Bach, many people are able to immediately identify other pieces by the same composer, and it seems that being part of a therapeutic team can teach more than many therapists' treatment manuals. More frequent alliance measures are needed, and as discussed above, they should perhaps be based on the concept of embedded alliance. A good alliance measure for a given therapy would include items related to this specific treatment feature, and research should be aimed at how to maximize implementation of these components (Hatcher, 2010, p. 16). This would be much more important than merging the CF into the manuals, which are already quite good. Reading the case formulations aloud in MBT groups is one such implementation of alliance in clinical practice. Such a research project is currently running at Oslo University Hospital. However, supplementing the manual with typical strategies is hereby called for.

We are evidently far from competing with the precision of quantum electrodynamics in the upper-right quadrant, but machine learning (one form of AI) is already suggested for alliance research and for predicting personalized process–outcome associations (Goldberg et al., 2020; Rubel et al., 2020). However, as computers are assumed to acquire an ability to mentalize within 15–20 years (MIT)—and perhaps inspired by recent reports of MBT seemingly surviving the current digitalization (e.g., therapy performed remotely due to COVID-19; Fonagy et al., 2020)—one might ask whether MBT done by computers with the ability to mentalize may be helpful for BPD patients in the future. At least, it seems that therapy done online is characterized by the same correlation between alliance and outcome as therapy performed face to face (Flückiger et al., 2018). This could be an interesting path to pursue in future research, as the results may indicate how much of the alliance effect is produced by physical presence (e.g., non-verbal communication). "Whatever aspects of the alliance are captured in Internet therapies, the alliance appears to relate to outcome, in a quantitative sense, similarly to face-to-face psychotherapy" (Flückiger et al., 2018, p. 332).

Importantly, while MBT should embrace the CFs, we should not stop researching the specific techniques.

> Crits-Christoph, Cooper, and Luborsky (1988) made a good beginning in showing that the accuracy of interpretation was associated with outcome in a psychodynamic treatment. (…) The paucity of information about technique–outcome associations might be attributable in part to editorial policies against publishing negative findings. (Benjamin & Critchfield, 2010, p. 124)

One possible path forward here is to rate different treatments with the MBT scales to get a better sense of how much specificity the scales have (Paper I gave an indication). However, there is a significant problem concerning rater bias (e.g., allegiance) here, and this is a kind of project ideally performed by less biased students. In terms of epistemic trust and pedagogic stance, one should perhaps gather more research data and investigate what expert therapists actually do (and how). One suggestion is to ask the patients what kind of strategies in the team/program they have been most influenced by. Additionally, the team should be aware of its administering skills and of the importance and impact of placebo. This process begins with faith in the treatment itself and in the method; for example, a significant portion of therapeutic change has been reported from the time of referral until the first session (Frank & Frank, 1993; Frank et al., 1963).

DBT (Linehan, 1993) is one example that mindfulness (e.g., Falkenström, 2010; Lilja et al., 2013) and other Eastern practices, such as Zen Buddhism, have been merged with evidence-based treatments for PDs. Therefore, there seems to be no reason not to include art therapy, body awareness, (Mahamudra) meditation, or even music therapy in treatment programs for PDs. Levy et al. (2010) identified two factors that contributed to difficulties in mentalization and interpersonal collaboration with BPD patients: 1) vacillating mental states and 2) deficits in executive attention. "Executive attention was related to therapeutic alliance, and this relationship was found to be mediated by in-session mental state vacillations" (Levy et al., 2010, p. 413). Executive attention, a top-down process involving ignoring extraneous stimuli, resolving cognitive conflict, and correcting errors, is disrupted in BPD by purely cognitive tasks (Berlin et al., 2005), although executive tasks with an affective component seem to more severely disrupt cognition for these patients (Silbersweig et al., 2007). Executive functioning can be altered through meditation (Tang & Posner, 2009) and skills-based treatments, such as DBT, "which focus on mindfulness training, and treatments geared toward developing greater mentalizing capacities by explicitly focusing on integrating alterations between mental states such as TFP may both affect the common neurocognitive correlate of executive attention but

by different routes" (Levy et al., 2010, p. 419). Importantly, MBT has implications for treatment, but as we have seen (e.g., mentalizing/minding the baby; e.g., Sadler et al. (2006) it also has implications for prevention. Prevention after all is the best treatment available, and our treatment models provide plenty of knowledge that should be made available for parents, teachers, and caretakers in our society. We are all part of the tea party, and it could be argued that our society is no stronger than the weakest link in the chain. Donald Trump, the 45th president of the United States, would perhaps disagree, but our united states of unconsciousness may also be important. Societies with large differences between people are shown to foster aggression, anxiety (e.g., Motesharrei et al., 2014; Wilkinson & Pickett, 2011), and perhaps even personality problems.

5.19   Strengths and limitations

The generalizability of the findings in Paper I is restricted by several limitations. In the current study, two to four treatment sessions per therapist were rated. We cannot exclude the possibility that repeated ratings of the same therapists may have artificially increased inter-rater conformity. Therefore, future studies should apply these scales to larger samples of both patients and therapists. Further, as the raters were not blinded to the treatment modality (PDG or MBT-G), the observed differences between therapies cannot exclusively be interpreted as reflecting the discriminant validity of the scale. Further, because of the limited number of videos analyzed, important further validity analyses, such as content, concurrent, criterion, and predictive validity, were not possible and must await a larger sample and set of analyses. As was the case in the study by Simonsen et al. (2019), the raters in Paper I knew most of the observed therapists, and our results could also be influenced by researcher allegiance. Simonsen et al. (2019) found that the sessions with high quality had the highest reliability. This could indicate that it is simpler to recognize prototypical MBT than to agree on the adherence and competence of interventions that fall outside or on the edge of the treatment model. However, they discovered that the ratings were influenced by the relationship between rater and therapist. Therefore, it is a strength that Paper I performed a G-study, disallowing such significant systematic error variance.

As discussed in Paper III, the study included two different versions of WAI. Therefore, differently formulated items may have affected the alliance ratings. However, even when given the same questionnaire, Falkenström et al. (2015) show that items in the WAI-SR can be interpreted differently due to language or cultural factors. As one could argue that the two

versions of the WAI, when recoded, are not more different than such cultural differences in interpretations of the items, one may find that this limitation is minor in nature (Paper III included both versions controlled for possible impacts). As we have seen above, in Elkin's (1999) concern about the effect of teams on psychotherapy studies we have a somewhat similar problem in terms of investigating team therapy, and one may perhaps argue that the results in Paper III should be interpreted as a product of the patients' alliance to the overall program. However, this would be in line with the philosophy of structured treatment programs for BPD. Further, as the initial ratings were high in Paper III, this may have caused a ceiling effect. However, it is a strength that Paper II and Paper III include measures of treatment fidelity and use LMMs to investigate alliance (Falkenström et al., 2013a). Another limitation to interpreting the association between alliance and outcome in Paper III, is that having an alliance with the group and epistemic trust in other patients seem to be crucial elements for successful healing for BPD patients (Antonsen, 2016). This dissertation, as well as Paper II and III, focused on the alliance in individual therapy. And investigated no alliance measure for the group component of the MBT program. Consequently, it is for future studies to investigate the impact of group alliance on outcome in MBT (and other BPD treatments).

It is a considerable strength that one of the authors in Paper II and III had no MBT allegiance, neither clinical experience in an MBT team nor specific MBT training (Erik Stänicke). The author of this dissertation is also a co-author of a new MBT manual (Karterud et al., 2020), which may have impacted certain cultural or political aspects of this thesis. For instance, one main bias of the current author and of one of the senior co-authors of Paper I and Paper II (SK) is that we find the original MBT model and what is performed in Norway to be slightly different in tune due to some differences in culture and approach to the length of training, and the target population. However, the new manual was written after all three papers in this study were finished, and we have tried our best to exhibit self-reflection in our interpretations (primarily relevant for Paper II).

By including both MBT-G sessions and psychodynamic groups (non-MBT sessions) in Paper I, we attempted to ensure sufficient variance to determine reliability. Low variance across treatments for several items indicates insubstantial differences between treatment modalities. One may wonder whether what is denoted MBT-G is more a disorder-specific competence in general (addressing the fact that the majority of the patients in this study had BPD). However, the CF approach would expect such overlaps to be substantial. "Cognitive-behavioral

therapies (CBT), psychodynamic therapy, experiential and existential therapies, as well as other psychological treatments, although apparently quite diverse, share much in common" (Imel & Wampold, 2008, p. 249). Therefore, Esperanto strategies for working with BPD patients are part of good MBT and are one reason why a CF approach to treatments of PDs (Bateman et al., 2018; Kernberg, 2016) seems fruitful.

At a conceptual level, there are several strengths and limitations when investigating the alliance construct in psychotherapy. For instance, it may seem confirming to the theory of the validity of the alliance that it is identified as the quintessential common factor (explaining more variance in outcome than the specific factors). "But correlation with outcome as a proof of construct validity is based on questionable logic" (Horvath, 2018), and there exists an impressive range of diversity among alliance measures. Consequently, the term alliance can point to a number of rather loosely related phenomena. Such a lack of consensus reduces the generalizability of the empirical findings and undermines scientific progress (Horvath, 2018). However, the variety in such instruments may also be recognized as a considerable strength, reflecting the fundamental importance of this aspect in psychotherapy. Mostly relevant for Paper II, the observed differences between client- and therapist-based self-reports of the quality of the alliance using similar or identical instruments likely reflect phenomenological perspectives on the relationship (Horvath, 2018). However, the observer-rated alliance seems to capture different aspects of the alliance than self-report-based measures do. Hence, "the phenomenological information the participants respond to and the observable data address different aspects of the relationship" (Horvath, 2018). This means that our interpretation of the therapeutic relationship would not necessarily be in tune with what the participants experienced, which would be the ideal in IPA.

It is a strength that the three applied methods in the current thesis have arguably measured MBT from three different quadrants or from all three of Habermas' (1986) validity claims— G-study: We/lower-left quadrant (consensus between raters); LMM: (Somewhat approaching) It/upper-right quadrant (statistical approach to data for 155 patients); and IPA: I/upper-left quadrant (transparent and justified inferences from purposeful sampling). It is also a strength that Paper III applied a linear multilevel model; Baldwin et al. (2007) "recommend that researchers use multilevel models or other innovative methods whenever they are studying variables that could conceivably vary among therapists, such as the alliance, treatment adherence, and treatment competence" (p. 851).

5.20   Conclusion

The current study includes three papers but has investigated four facets of MBT and has hopefully taken a few steps towards filling in four missing pieces in the larger jigsaw puzzle of measuring MBT. The group component now has a reliable fidelity measure, already implemented and applied in research. In addition, temporal development of the working alliance is shown to characterize favorable outcomes in MBT. Bridging the CF and EST approaches, this dissertation argues that manuals and evidence-based treatments (for BPD) bring important value and that such programs could be seen as overarching strategies to foster a strong and long-lasting working alliance (e.g., with particular emphasis on tasks and goals; Paper II and Paper III). Competence in particular CFs, such as forming a working alliance across a range of patients or having a high level of facilitative interpersonal skills, generally predicts more variance in outcome than competence (which presupposes adherence) in delivering specific treatment potions (Wampold & Imel, 2015). In Paper II, I argued that the quality/competence ratings in MBT mainly measure competence in particular CFs, such as strategic competence and working alliance. Consequently, "[s]killful application of MBT includes an overarching ability to navigate not defined by the MBT manuals" (Paper II, p. 144), that is, a strategic competence. Therefore, future research should investigate such overarching strategies and pedagogic stance (as pedagogic interventions color delivered treatments, typically "camouflaged" as Item 16). This also resonates with the recent focus on epistemic trust in MBT. The quality (competence) of MBT seems to largely overlap with the working alliance construct introduced by Bordin in 1979. The main conclusion of this thesis is in line with an unexplored field in psychotherapy research proposed by Hatcher (2010), who argued that a good alliance measure for a given therapy would include items related to this specific treatment feature. Consequently, I suggest that future measurement tools should focus on embedded alliance. Clinical supervision adheres to such a conclusion.

As some therapists are better at producing a good working alliance no matter what method they employ for unknown reasons (Lemma et al., 2011), there was a need to investigate how skilled therapists fostered the therapeutic alliance. Manualized treatments can be viewed as attempts to provide clinicians with some guidelines from expert therapists on what kinds of strategies or interventions are considered to be helpful or to nurture the alliance for certain patients. However, the manual (treatment approach) must be adapted to the specific patient,

and the merger of the working alliance with the specific technique(s) in the term embedded alliance seems like a fruitful path to pursue in future research and in clinical practice.

It has been claimed that despite there being many theories about what brings about change, we know little of how change actually occurs. Our study adds but a small atom of insight, but it still seems possible to establish a working alliance even with severely disturbed patients (paranoid PD and BPD) and that therapists who succeed in this succeed in battling the patient's comfort zone in a way that conveys a strong belief in the practiced method and that generates epistemic trust. The bond appeared to be an important asset in this process, and it seemed to develop through an increased epistemic trust generated by repeated challenges, which resulted in an improved reality for the patient previously. As predicted by the CF approach, Paper II and Paper III signaled that liking the therapist (positive bond) was less predictive for change; that is, a focus on tasks and goals is the core of the therapy. Therapists who are brought out of their own comfort zone seemed to lose track of the therapeutic project, and it seems that providing therapists with examples of how to handle typical countertransference reactions and resistance by the patient will be an important next step. The same goes for interventions targeting epistemic trust and probably pedagogic interventions. Psychotherapeutic potions may not be exclusively for talented therapists, and deliberate practice is suggested as key to integrate techniques to foster alliance in one's own practice.

> Provided that the person training to become a therapist is not himself severely disturbed, practical clinical training and competent supervision will make it possible for him to internalize these qualities and integrate them as parts of his personality. As we become better in articulating these common factors—or in making some of the nonspecific factors more specific—parts of this learning and internalization process can be accelerated. (Jørgensen, 2004, p. 536)

As all Euclideans know, the circle is hard to square (despite our four quadrants), and certainly prefers completion. Therefore, I will let Habermas (1986) conclude our attempt to measure MBT:

> When valid, [general] theories hold for all who can adopt the position of the inquiring subject. When valid, general interpretations hold for the inquiring subject and all who can adopt its position only to the degree that those who are made the object of individual interpretations know and recognize themselves in these interpretations. (pp. 261–2)

Tables & Figures

Table 1: Frequency/number (adherence) of specific MBT interventions in sessions with various overall ratings (1–2; 3–5; 6–7)

MBT Item	Average MBT	MBT sessions rated 6–7	MBT sessions rated 3–5	MBT sessions rated 1–2
2. Exploration, curiosity, and a not-knowing stance	12.4	16	14	3
3. Challenging	1.6	2.5	1	2
7. Acknowledging positive mentalizing	1.9	5	0.5	1
9. Psychic equivalence	0.4	1.5	0	0
10. Affect focus	9.5	11.5	9	8
11. Affect and interpersonal events	4.0	3	5	3
12. Stop and rewind	0.4	0.5	0	1
13. Validating feelings	4.3	9.5	2.5	1
14. Relation to therapist	5.3	11	2	5
15. Counter-transference	1.4	1.5	1.5	1
16. Validating understanding	19.9	21.5	23	10
17. Integrating group experiences	1.6	2.5	0.5	3
Number of interventions	62.8	86	59	38
N (number of MBT sessions)	327	97	164	66

Table 2: The nine prototypical versions of interventions targeting impaired epistemic trust (including missed opportunities)

1: Mistaking or confusing the recipe, rulebook, or "correct view" with reality
2: Mistaking or confusing others' dependency, gratitude, or relational valence with reality
3: Mistaking or confusing achievements, actions, or superficial mirroring (e.g., looks, status, clothes) with reality
4: Mistaking or confusing one's own inner world (typically strong emotions or looking for something that feels "right") with reality
5: Mistaking or confusing knowledge and knowledge about knowledge with reality
6: Mistaking or confusing idealization of life project with reality
7: Mistaking or confusing (endless) possibilities and fantasies with reality
8: Mistaking or confusing strong passion or intense pain with reality
9: Mistaking or confusing love or coziness with reality

Table 3: Item descriptives, G-study results, and D-study results for different measurement designs

	Adherence/frequency				Competence/quality			
Item	Mean[a]	SD[a]	G rel.	G abs.	Mean[a]	SD[a]	G rel.	G abs.
1. Boundaries	6.18	3.34	0.89	0.86	4.13	1.00	0.90	0.89
2. Phases	3.98	3.60	0.93	0.90	3.34	1.85	0.96	0.96
3. Turn taking	5.48	4.61	0.95	0.95	3.35	1.86	0.95	0.95
4. External events	5.38	3.44	0.94	0.92	3.38	1.44	0.93	0.92
5. Events in group	3.31	2.78	0.85	0.85	3.09	1.42	0.90	0.87
6. Care for group					4.36	1.23	0.91	0.91
7. Authority					4.33	1.42	0.93	0.93
8. Group norms	2.36	2.64	0.83	0.83	2.81	1.88	0.78	0.78
9. Cooperation	1.53	1.93	0.86	0.84	2.10	1.71	0.96	0.95
10. Warmth					4.50	1.15	0.92	0.91
11. Exploration	16.03	6.79	0.87	0.80	3.98	1.32	0.86	0.86
12. Unwarranted beliefs	2.48	2.64	0.88	0.88	3.00	1.35	0.84	0.82
13. Emotional arousal					3.68	1.29	0.93	0.93
14. Acknowledging	1.64	1.84	0.77	0.77	2.69	1.56	0.86	0.85
15. Pretend mode					2.23	1.58	0.67	0.64
16. Psychic equivalence	1.35	1.70	0.63	0.63	2.64	1.77	0.86	0.86
17. Affect focus	14.85	7.03	0.91	0.85	4.28	1.71	0.93	0.93
18. Stop and rewind	0.65	1.08	0.88	0.88	1.80	1.81	0.85	0.84
19. Relationship	5.00	5.29	0.94	0.93	3.30	1.63	0.88	0.88
Overall rating	3.76	1.76	0.97	0.97	3.80	1.67	0.96	0.96
Mean values	4.93	2.69	0.87	0.86	3.34	1.27	0.89	0.88

a) Grand mean and standard deviation of scores across raters and sessions
b) This scale is not rated for adherence, only quality.

Table 4: Mean ratings of PDG and MBT-G

Item	Adherence PDG	Adherence MBT-G	Competence PDG	Competence MBT-G
1. Boundaries	5	8	4	5
2. Phases	1	7	2	5
3. Turn taking	2	9	2	5
4. External events	5	5	3	4
5. Events in the group	3	4	3	4
6. Care for group	*Not rated*		4	5
7. Authority	*Not rated*		4	5
8. Group norms	6	5	3	3
9. Cooperation	4	0	2	2
10. Warmth	*Not rated*		4	5
11. Exploration	14	5	3	5
12. Unwarranted beliefs	1	1	3	3
13. Emotional arousal	*Not rated*		3	4
14. Acknowledging	3	4	2	3
15. Pretend mode	*Not rated*		2	2
16. Psychic equivalence	0	1	2	3
17. Affect focus	12	14	3	5
18. Stop and rewind	0	1	1	3
19. Relationship	9	7	3	4
Overall	5	6	3	5

Table 5: Sources of variation for items with high versus low reliability on quality with 5R: Percentages of total variation

Item	T	R	S:T	TR	RS:T	Abs G
15. Pretend modus	17.1	7.1	9.3	12.6	53.9	0.64
8. Group norms	0	1.8	41.2	0	57	0.78
12. Unwarranted beliefs	24.5	7.3	23	11.2	34	0.82
18. Stop and rewind	22.1	4.2	28.2	0	45.5	0.84
14. Acknowledging	50.6	2.6	2.4	0	44.4	0.85
11. Exploration	33.6	1.8	21	10.4	33.2	0.86
16. Psychic equivalence	20.5	1.1	33.8	4.3	40.3	0.86
5. Events in the group	42.4	10.7	14.7	3.7	28.5	0.87
19. Relationship	0	0.4	59.3	0	40.3	0.88
1. Boundaries	34.3	6.2	26.4	5.5	27.6	0.89
6. Care for group	41.3	1.6	25	2.6	29.6	0.91
10. Warmth	58.1	0.8	9.9	0	31.3	0.91
4. External events	46.5	3.5	24.5	0	25.6	0.92
7. Authority	54.2	1.6	18.3	12.1	13.7	0.93
13. Emotional arousal	43.1	0	28.6	5.8	22.4	0.93
17. Affect focus	44.6	0	27.5	0	27.9	0.93
3. Turn taking	62.7	0.8	16.7	5.9	14	0.95
9. Cooperation	0	4.7	79.9	0	15.4	0.95
2. Phases	65.6	0	16.4	0	18	0.96
Overall	61.7	0	21.2	0	17.1	0.96

T: Between-therapist variation
R: Variation in how much raters observe
S:T: Therapist variation across sessions
TR: Variation in raters' ranking of therapists
RS:T: Residual (including error) variance
Abs G: Agreement on exact scores

Table 6: Adherence ratings on MBT-I-ACS

Item	Session A	Session B	Session C	Session D
2. Exploration, curiosity, and a not-knowing stance	**39**	**31**	**6**	**14**
3. Challenging unwarranted beliefs	**8**	**5**	**0**	**0**
7. Acknowledging positive mentalizing	**0**	**0**	**0**	**0**
9. Psychic equivalence	**1**	**2**	**3**	**0**
10. Affect focus	**16**	**20**	**6**	**6**
11. Affect and interpersonal events	**14**	**14**	**3**	**4**
12. Stop and rewind	**0**	**0**	**0**	**0**
13. Validation of emotional reactions	**12**	**4**	**0**	**4**
14. Transference and the relation to the therapist	**14**	**8**	**12**	**4**
15. Use of countertransference	**5**	**0**	**0**	**1**
16. Monitoring own understanding and correcting misunderstanding	**21**	**41**	**20**	**13**
17. Integrating experiences from concurrent group therapy	**30**	**68**	**0**	**6**
Overall score for entire session	**6**	**6**	**2**	**2**

Table 7: Competence ratings on MBT-I-ACS

Item	Session A	Session B	Session C	Session D
1. Engagement, interest, and warmth	7	6	3	3
2. Exploration, curiosity, and a not-knowing stance	7	6	2	2
3. Challenging unwarranted beliefs	6	5	2	0
4. Adaptation to mentalizing capacity	6	6	2	2
5. Regulation of arousal	6	6	2	2
6. Stimulating mentalization through the process	7	6	2	2
7. Acknowledging positive mentalizing	0	0	2	2
8. Pretend mode	0	5	2	2
9. Psychic equivalence	6	5	2	0
10. Affect focus	5	6	3	2
11. Affect and interpersonal events	6	5	2	2
12. Stop and rewind	0	0	2	0
13. Validation of emotional reactions	5	4	2	4
14. Transference and the relation to the therapist	4	6	2	2
15. Use of countertransference	6	0	2	2
16. Monitoring own understanding and correcting misunderstanding	6	6	3	2
17. Integrating experiences from concurrent group therapy	7	7	2	2
Overall score for entire session	6	6	2	2

Table 8: Total number of interventions and interventions rated as MBT or not MBT

	Session A	Session B	Session C	Session D
Interventions	165	180	111	276
Not MBT	25	44	69	224
MBT	140	136	42	52
MBT ratings[1]	160	193	50	52
Index	85%	76%	38%	19%

[1] One intervention can have multiple MBT ratings. For example, in session A, there were 20 MBT interventions with more than one adherence rating.

Table 9: Transcript from challenging segment in session C with comments and indications of MBT-I-ACS ratings

	Verbatim material (translated from Norwegian)	Comments (item number)
Patient	**No, I didn't. More that it seemed like it was my… my poor self confidence, or my… well, my experience of how people are at University. And… that you maybe wasn't open to see that there are more people seeing things that way… or more people experiencing it that way. That it's not just a thing that I experience because I am the way that I am.**	**Patient tries to mentalize and gives an account for the view.**
Therapist	**Yes. Eh… Mhm.**	
Patient	**That you tried… Well, it seemed like you were interpreting it as my individual experience, and not something to do with them. (Clears throat.)**	**Patient tries to mentalize the therapist.**
Therapist	**Mm… eh… you might be right about that, that I did it that way. Yes. Because… I think that… Because… I think like this, that because… you were saying that… because you got irritated when I said that.**	**14, 10. Does not facilitate mentalizing.**
Patient	**Angry.**	**Patient corrects**
Therapist	**Angry (laughs a little).**	**10**
Patient	**Don't use the word irritated when I'm not irritated.**	**Patient corrects**

Table 10: Transcript from session D with comments and indications of MBT-I-ACS ratings

	Verbatim material (translated from Norwegian)	Comments (item number)
Patient	**I have been so tired.**	
Therapist	**Yes. I understand you well, P, and I have been thinking that this… this has been a… really an intense experience for you.**	**Guesstimates**
Patient	**Mm.**	
Therapist	**Hard… and that maybe… that what… I don't know if you think like this, but I think that it's got… that what you are in the middle of now has to do with that… that… rape**	**Guesstimates**
Patient	**Mm.**	
Therapist	**What you… do you think like that too? That that is what…**	
Patient	**Yes, I think that too.**	

Table 11: Transcript from session D with comments and indications of MBT-I-ACS ratings

	Verbatim material (translated from Norwegian)	Comments (item number)
Patient	I think that after Christmas and New Year it will be a new beginning.	
Therapist	Yes.	
Patient	And I will try to sit for an exam now.	
Therapist	Yes. Quite some… quite a lot going on for you this autumn, I think.	
Patient	Yes, it's not supposed to be easy	
Therapist	No.	

Table 12: Transcript from challenging segment in session C with comments and indications of MBT-I-ACS ratings

	Verbatim material (translated from Norwegian)	Comments (item number)
Patient	Yes. Did you send the statement?	
Therapist	I did at least send… Let me check if… If I'm sure that I've done it.	
Patient	Yes. Because it should have been sent. I think so… I have at least told them that it is on its way, so that…	
Therapist	Yes.	
Patient	… now that you have postponed that deadline and all…yesterday.	
Therapist	The deadline.. No, the time is up… it has been sent.	
Patient	it is up, yes. Yes.	
Therapist	Yes.	
Patient	Then it's ok.	
Therapist	So it is… mm.	
Patient	Yes. I got irritated with you last session, didn't I, I got angry with you. You said I was irritated, but I wasn't, I was angry.	
Therapist	You were angry with me… yes… yes.. mm.	14, 10

Table 13: Transcript from challenging segment in session A with comments and indications of MBT-I-ACS ratings

	Verbatim material (translated from Swedish)	Comments (item number)
Therapist	**It will be difficult, doesn't?**	**10, 17**
Patient	**Yes.**	
Therapist	**The only way to find out is to go there.**	**16, 17**
Patient	**Yes.**	
Therapist	**But then of course you think I'm a fool telling you this for the hundredth time.**	**17**
Patient	**Yes.**	
Therapist	**At the same time, I think like this: Now that we're talking about it, I try in a way, well … it... it is quite difficult, because I can't hide that I think that's good for you to go there. Just because I happen to think so?! But at the same time, I feel that I nag you about this a lot. And then I think like this: Is it because I nag on you, that you say yes, that you want to go there, because you don't go there. And then I feel...well, what am I doing..... and I feel disappointed in a way. We talk about it and you say you will go there and then you don't....**	**14, 15, 17 Confronts at the maximum of what the patient can tolerate. Lets his own mind be transparent and uses his own feelings.**
Patient	**Over and over again. Over and over again.**	
Therapist	**Yes**	
Patient	**Yes. I understand. I understand. I understand.**	
Therapist	**I feel something too.**	**15**

Figure 1: Wilber's four quadrants (I, We, It, Its) incorporate Habermas' three validity claims (I, We, It) in his unifying theory, which forms the epistemological basis for the current thesis.
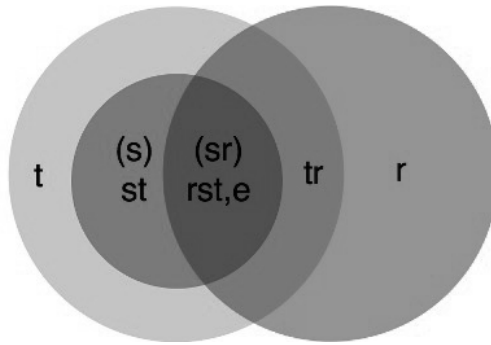
Figure 2: Venn diagram of the variance components in the (s:t) x r design

Figures 3, 4, 5, and 6 display the mean intervention profiles (ratings on all items) for quality and adherence for MBT-G (2014) and PD (2006) groups
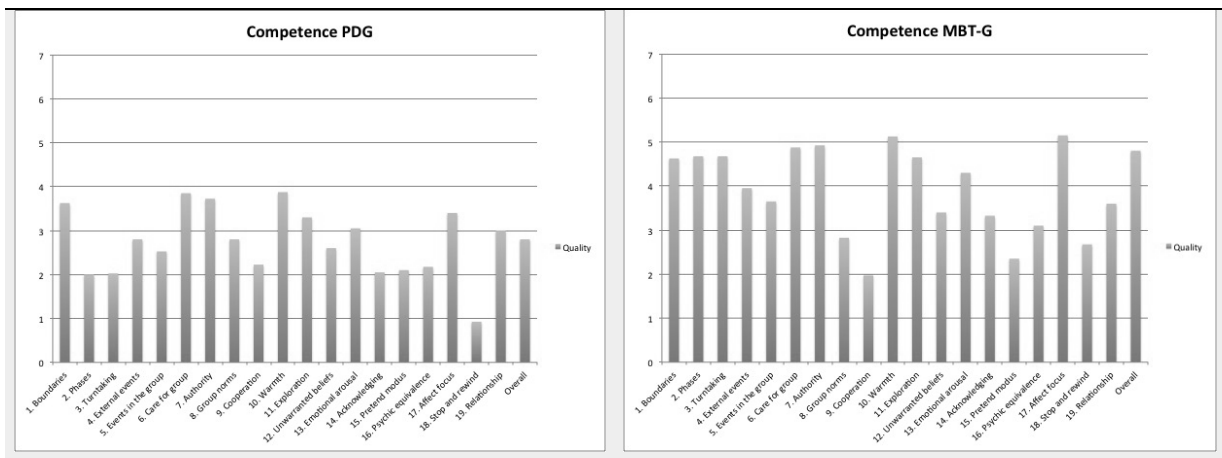
Figures 3 and 4: The mean intervention profiles for competence scores for groups from for PDG and MBT-G, respectively
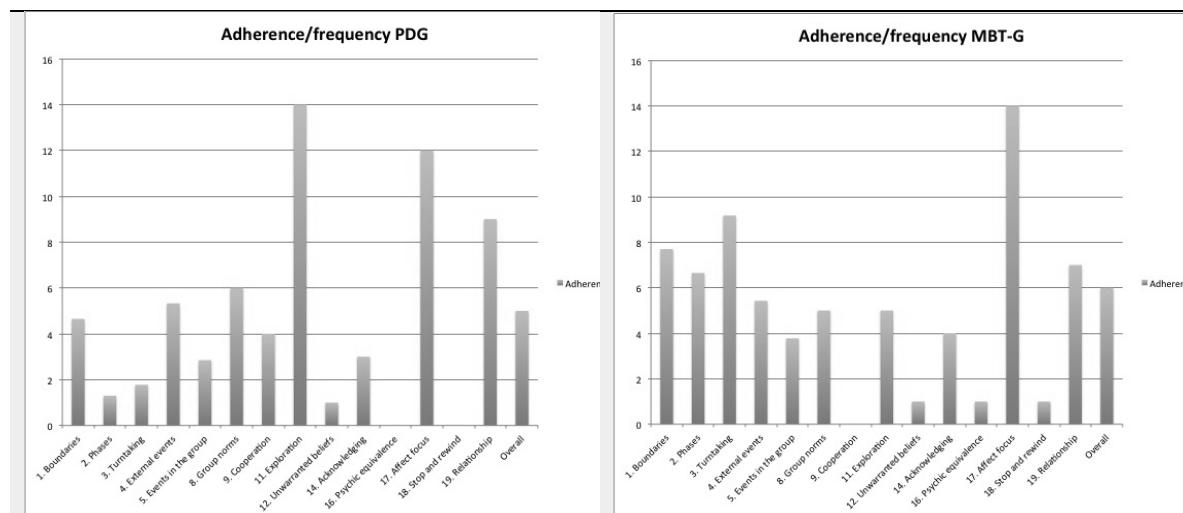
Figures 5 and 6: The mean intervention profiles for adherence/frequency scores for PDG and MBT-G, respectively

MBT Index – Good MBT		MBT Index – Poor MBT	
MBT	80%	MBT	28%
Not MBT	20%	Not MBT	72%
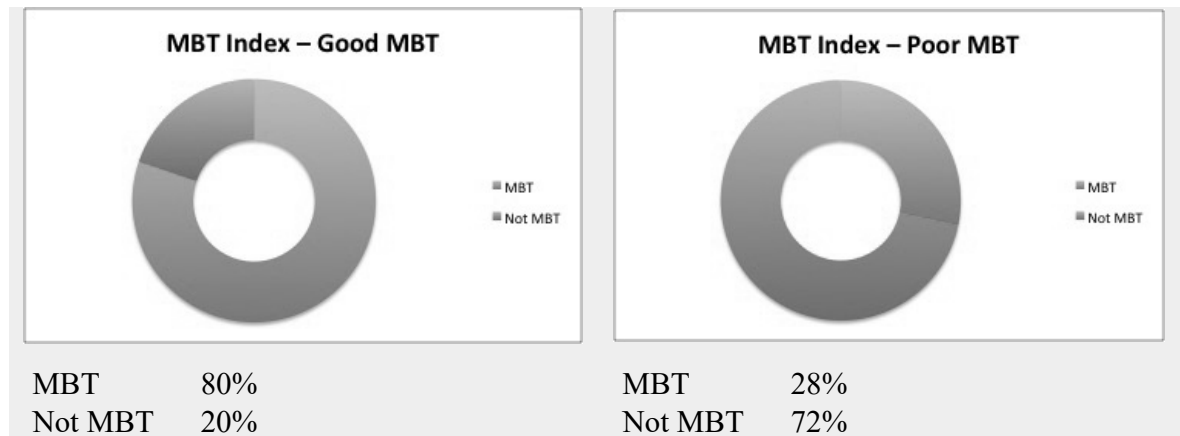
Figure 7: Percentage of interventions rated for adherence to MBT in good and poor sessions (good sessions: A and B; poor sessions: C and D): Index of MBT interventions.

Figure 8: Mean MBT rating for good versus poor sessions (good sessions: A and B; poor sessions: C and D).

Appendices

Appendix A. The MBT-I-ACS with anchor points (level 4 rating). Adherence and competence are rated on a scale from 0 to 7, and the rater's basic assumption when grading competence should be that the therapist is average (score of 4).

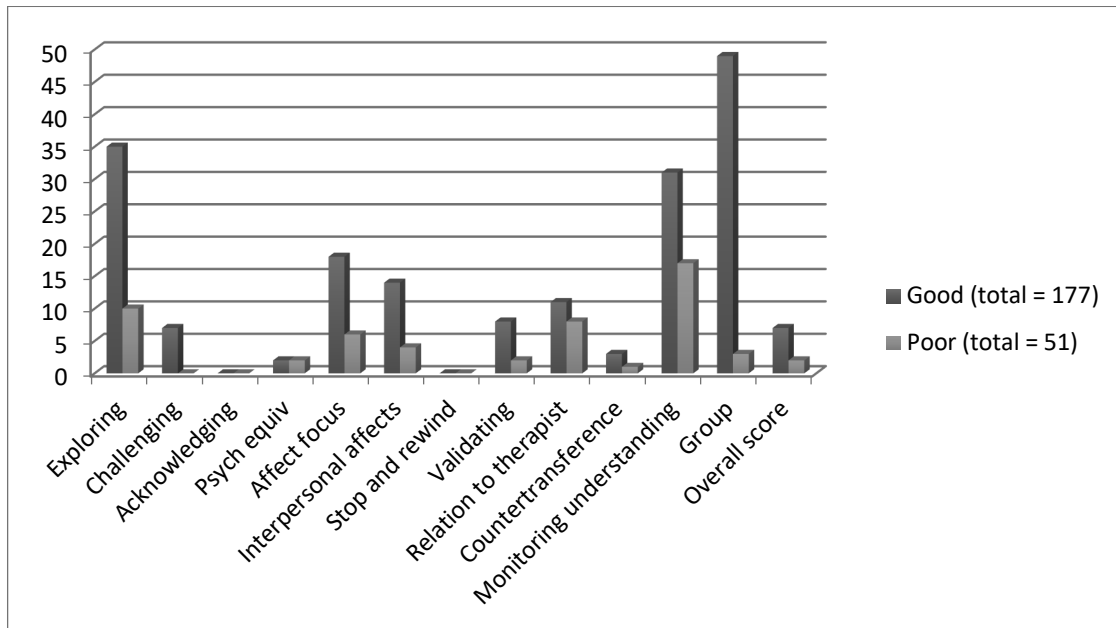| Item # and name	Adherence rating	Notes for "good enough" quality
1. Engagement, interest and warmth	This item is not rated for adherence	4: The therapist appears genuinely warm and interested. The rater gets the impression that the therapist care. Several concrete comments communicate this positive attitude
2. Exploration, curiosity and a not-knowing stance		4: The therapist poses appropriate questions designed to promote exploration of the patient's and others mental states, motives and affects and communicate a genuine interest in finding out more about them
3. Challenging unwarranted beliefs		4: The therapist confronts and challenges unwarranted opinions about oneself or others in an appropriate manner
4. Adaptation to mentalizing capacity	This item is not rated for adherence	4: The therapist seems to have adapted to the patient's mentalizing level and the interventions are for the most part short, concise and unpretentious
5. Regulation of arousal	This item is not rated for adherence	4: The therapist plays an active role in terms of maintaining emotional arousal at an optimal level (not too high so that the patient looses his or her ability to mentalize; not too low so that the session becomes meaningless emotionally)
6. Stimulating mentalization through the process	This item is not rated for adherence	4: The aim of the interventions clearly seems to be to stimulate the mentalizing of experiences of self and others in an ongoing process and is less concerned about content and interpretation of content in order to promote insight
7. Acknowledging positive mentalizing		4: The therapist identifies and explores good mentalization and this is accompanied by approving words or judicious praise
8. Pretend mode	This item is not rated for adherence	4: The therapist identifies pretend mode and intervenes to improve mentalizing capacity
9. Psychic equivalence		4: The therapist identifies psychic equivalence functioning and intervenes to improve mentalizing capacity
10. Affect focus		4: The interventions focus primarily on affects, more than on behavior. The attention is directed at affects as they are expressed in the here and now, and particularly in terms of the relationship between patient and therapist
11. Affect and interpersonal events		4: The therapist connects emotions and feelings to recent or immediate interpersonal events
12. Stop and rewind		4: The therapist identifies at least one incident in which the patient reacts in a maladaptive way to an interpersonal event, then tries to slow down the pace and find out about the incident step by step
13. Validation of emotional reactions		4: The therapist expresses a normative view on the warranted nature of the patient's emotional reaction(s) after these are sufficiently investigated and understood
14. Transference and the relation to the therapist		4: The therapist comments on and attempts to explore – together with the patient -- how the patient relates to the therapist during the session and stimulates reflections on alternative perspectives whenever appropriate
15. Use of countertransference		4: The therapist actively utilizes his/her own feelings and thoughts about the relationship to the patient and attempts by this to stimulate an exploration of the relationship between them
16. Monitoring own understanding and correcting misunderstanding		4: The therapist checks out his/her understanding of the patient's state of mind and to what extent this corresponds with the patient's understanding. Then he/she lets his/her own understanding be influenced by the patient's understanding and openly admits to any misunderstanding whenever they occur
17. Integrating experiences from		4: The therapist stimulates exploration of the patient's experiences from the group therapy sessions and helps to integrate the material so that the treatment as a whole is coherent

concurrent group therapy		

References

Aitchison, J. (2000). *The seeds of speech: Language origin and evolution*. Cambridge University Press.

Alden, L. E., Wiggins, J. S., & Pincus, A. L. (1990). Construction of circumplex scales for the Inventory of Interpersonal Problems. *Journal of personality assessment, 55*(3-4), 521-536.

Allen, J. G. (2012). *Restoring mentalizing in attachment relationships: Treating trauma with plain old therapy*. American Psychiatric Pub.

Allen, J. G., Fonagy, P., & Bateman, A. W. (2008). *Mentalizing in clinical practice*. American Psychiatric Pub.

Allison, E., & Fonagy, P. (2016). When is truth relevant? *The Psychoanalytic Quarterly, 85*(2), 275-303.

Antonsen, B. T. (2016). *Long-term clinical outcome of psychotherapeutic treatment for patients with personality disorders: findings from a randomized study* [Doctoral Thesis, University of Oslo, Norway].

Antonsen, B. T., Kvarstein, E. H., Urnes, Ø., Hummelen, B., Karterud, S., & Wilberg, T. (2017). Favourable outcome of long-term combined psychotherapy for patients with borderline personality disorder: six-year follow-up of a randomized study. *Psychotherapy research, 27*(1), 51-63.

Arnevik, E., Wilberg, T., Monsen, J. T., Andrea, H., & Karterud, S. (2009a). A cross-national validity study of the Severity Indices of Personality Problems (SIPP-118). *Personality and mental health, 3*(1), 41-55.

Arnevik, E., Wilberg, T., Urnes, Ø., Johansen, M., Monsen, J. T., & Karterud, S. (2009b). Psychotherapy for personality disorders: Short-term day hospital psychotherapy versus outpatient individual therapy–a randomized controlled study. *European Psychiatry, 24*(2), 71-78.

Bailey, R., & Ogles, B. (2019). Common Factors as a Therapeutic Approach: What Is Required? *Practice Innovations, 4*(4), 241-254.

Baker, T., & McFall, R. (2014). The promise of science-based training and application in psychological clinical science. *Psychotherapy, 51*(4), 482-486.

Baldwin, S. A., & Imel, Z. (2013). Therapist effects: Findings and methods. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6 ed., pp. 258–297).

Baldwin, S. A., Wampold, B. E., & Imel, Z. E. (2007). Untangling the alliance-outcome correlation: Exploring the relative importance of therapist and patient variability in the alliance. *Journal of Consulting and Clinical Psychology, 75*(6), 842.

Bales, D., van Beek, N., Smits, M., Willemsen, S., Busschbach, J. J., Verheul, R., & Andrea, H. (2012). Treatment outcome of 18-month, day hospital mentalization-based treatment (MBT) in patients with severe borderline personality disorder in the Netherlands. *Journal of personality disorders, 26*(4), 568–582.

Bales, D. L., Timman, R., Luyten, P., Busschbach, J., Verheul, R., & Hutsebaut, J. (2017a). Implementation of evidence-based treatments for borderline personality disorder: The impact of organizational changes on treatment outcome of mentalization-based treatment. *Personality and mental health, 11*(4), 266–277.

Bales, D. L., Verheul, R., & Hutsebaut, J. (2017b). Barriers and facilitators to the implementation of mentalization-based treatment (MBT) for borderline personality disorder. *Personality and mental health, 11*(2), 118-131.

Baltzersen, Å.-L. (2021). Moving forward: closing the gap between research and practice for young people with BPD. *Current Opinion in Psychology, 37*, 77–81.

Barber, J. P. (2009). Toward a working through of some core conflicts in psychotherapy research. *Psychotherapy research, 19*(1), 1-12.

Barber, J. P., Connolly, M. B., Crits-Christoph, P., Gladis, L., & Siqueland, L. (2000). Alliance predicts patients' outcome beyond in-treatment change in symptoms. *Journal of Consulting and Clinical Psychology, 68*(6), 1027–1032.

Barber, J. P., Crits-Christoph, P., & Luborsky, L. (1996). Effects of therapist adherence and competence on patient outcome in brief dynamic therapy. *Journal of Consulting and Clinical Psychology, 64*(3), 619.

Barber, J. P., Gallop, R., Crits-Christoph, P., Frank, A., Thase, M. E., Weiss, R. D., & Beth Connolly Gibbons, M. (2006). The role of therapist adherence, therapist competence, and alliance in predicting outcome of individual drug counseling: Results from the National Institute Drug Abuse Collaborative Cocaine Treatment Study. *Psychotherapy research, 16*(02), 229–240.

Barber, J. P., Khalsa, S.-R., & Sharpless, B. A. (2010). The validity of the alliance as a predictor of psychotherapy outcome. In J. C. Muran & J. P. Barber (Eds.), *The therapeutic alliance: An evidence-based guide to practice* (pp. 29–43). The Guilford Press.

Barber, J. P., Liese, B. S., & Abrams, M. J. (2003). Development of the cognitive therapy adherence and competence scale. *Psychotherapy research, 13*(2), 205-221.

Barber, J. P., Triffleman, E., & Marmar, C. (2007). Considerations in treatment integrity: Implications and recommendations for PTSD research. *Journal of Traumatic Stress, 20*(5), 793–805.

Barlow, D. H. (2004). Psychological treatments. *American Psychologist, 59*(9), 869.

Barlow, D. H., Andrasik, F., & Hersen, M. (2007). *Single case experimental designs*. Allyn and Bacon London.

Barnicot, K., Katsakou, C., Bhatti, N., Savill, M., Fearns, N., & Priebe, S. (2012). Factors predicting the outcome of psychotherapy for borderline personality disorder: a systematic review. *Clinical Psychology Review, 32*(5), 400-412.

Barnicot, K., Katsakou, C., Marougka, S., & Priebe, S. (2011). Treatment completion in psychotherapy for borderline personality disorder–a systematic review and meta-analysis. *Acta Psychiatrica Scandinavica, 123*(5), 327-338.

Bateman, A., Bales, D., & Hutsebaut, J. (2012). A quality manual for MBT. Draft presented at Nordic MBT-network meeting, Oslo,

Bateman, A., Campbell, C., Luyten, P., & Fonagy, P. (2018). A mentalization-based approach to common factors in the treatment of borderline personality disorder. *Current Opinion in Psychology, 21*, 44-49.

Bateman, A., Constantinou, M. P., Fonagy, P., & Holzer, S. (2020). Eight-year prospective follow-up of mentalization-based treatment versus structured clinical management for people with borderline personality disorder. *Personality Disorders: Theory, Research, and Treatment*, Advance online publication.

Bateman, A., & Fonagy, P. (1999). Effectiveness of partial hospitalization in the treatment of borderline personality disorder: a randomized controlled trial. *American Journal of Psychiatry, 156*(10), 1563-1569.

Bateman, A., & Fonagy, P. (2001). Treatment of borderline personality disorder with psychoanalytically oriented partial hospitalization: an 18-month follow-up. *American Journal of Psychiatry, 158*(1), 36-42.

Bateman, A., & Fonagy, P. (2006). Mentalizing and borderline personality disorder. In J. G. Allen & P. Fonagy (Eds.), *Handbook of mentalization-based treatment* (pp. 185–200). John Wiley & Sons, Hoboken, NJ.

Bateman, A., & Fonagy, P. (2009). Randomized controlled trial of outpatient mentalization-based treatment versus structured clinical management for borderline personality disorder. *American Journal of Psychiatry, 166*(12), 1355-1364.

Bateman, A., & Fonagy, P. (2013). Impact of clinical severity on outcomes of mentalisation-based treatment for borderline personality disorder. *The British Journal of Psychiatry, 203*(3), 221-227.

Bateman, A., & Fonagy, P. (2016). *Mentalization-based treatment for personality disorders: A practical guide*. Oxford University Press.

Bateman, A. W. (2007). Psychotherapy in psychiatry: Integration and assimilation. *International Review of Psychiatry, 19*(1), 1–4.

Beck, A. T., Davis, D. D., & Freeman, A. (2015). *Cognitive therapy of personality disorders*. Guilford Publications.

Beck, E., Bo, S., Jørgensen, M. S., Gondan, M., Poulsen, S., Storebø, O. J., Fjellerad Andersen, C., Folmo, E., Sharp, C., & Pedersen, J. (2020). Mentalization-based treatment in groups for adolescents with borderline personality disorder: a randomized controlled trial. *Journal of Child Psychology and Psychiatry, 61*(5), 594-604.

Bein, E., Anderson, T., Strupp, H., Henry, W., Schacht, T., Binder, J., & Butler, S. (2000). The effects of training in time-limited dynamic psychotherapy: Changes in therapeutic outcome. *Psychotherapy research, 10*(2), 119-132.

Benjamin, L. S., & Critchfield, K. L. (2010). An interpersonal perspective on therapy alliances and techniques. In J. C. Muran & J. P. Barber (Eds.), *The therapeutic alliance: An evidence-based guide to practice* (pp. 123-149). Guilford Press.

Benjamin, L. S., & Karpiak, C. P. (2001). Personality disorders. *Psychotherapy: Theory, Research, Practice, Training, 38*(4), 487–491.

Benton, T., & Craib, I. (2010). *Philosophy of social science: The philosophical foundations of social thought*. Macmillan International Higher Education.

Berlin, H. A., Rolls, E. T., & Iversen, S. D. (2005). Borderline personality disorder, impulsivity, and the orbitofrontal cortex. *American Journal of Psychiatry, 162*(12), 2360-2373.

Beutler, L. E. (1999). Manualizing flexibility: The training of eclectic therapists. *Journal of clinical psychology, 55*(4), 399-404.

Beutler, L. E., Malik, M., Alimohamed, S., Harwood, T. M., Talebi, H., Noble, S., et al. (2004). Therapist variables. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (pp. 227−306). Wiley.

Bhatia, S. C., Madabushi, J., Kolli, V., Bhatia, S. K., & Madaan, V. (2013). The Bhagavad Gita and contemporary psychotherapies. *Indian journal of psychiatry, 55*(Suppl 2), S315.

Binder, J. L., & Henry, W. P. (2010). Developing skills in managing negative process. In J. C. Muran & J. P. Barber (Eds.), *The therapeutic alliance: An evidence-based guide to practice* (pp. 285–304). Guilford.

Bion, W. R. (1977). *Seven servants: Four works*. Jason Aronson.

Blatt, S. J., Zuroff, D. C., Quinlan, D. M., & Pilkonis, P. A. (1996). Interpersonal factors in brief treatment of depression: Further analyses of the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Journal of Consulting and Clinical Psychology, 64*(1), 162.

Bloom, H. (2014). *The western canon: The books and school of the ages*. Houghton Mifflin Harcourt.

Bo, S., Sharp, C., Fonagy, P., & Kongerslev, M. (2017). Hypermentalizing, attachment, and epistemic trust in adolescent BPD: Clinical illustrations. *Personality Disorders: Theory, Research, and Treatment, 8*(2), 172.

Bohart, A. C. (2000). The client is the most important common factor: Clients' self-healing capacities and psychotherapy. *Journal of Psychotherapy Integration, 10*(2), 127-149.

Boldrini, T., Nazzaro, M. P., Damiani, R., Genova, F., Gazzillo, F., & Lingiardi, V. (2018). Mentalization as a predictor of psychoanalytic outcome: An empirical study of transcribed psychoanalytic sessions through the lenses of a computerized text analysis measure of reflective functioning. *Psychoanalytic psychology, 35*(2), 196.

Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, research & practice, 16*(3), 252.

Bordin, E. S. (1983). A working alliance based model of supervision. *The Counseling Psychologist, 11*(1), 35-42.

Bordin, E. S. (1994). Theory and research on the therapeutic working alliance: New directions. In A. O. Horvath & L. S. Greenberg (Eds.), *The working alliance: Theory, research, and practice* (pp. 13–37). Wiley.

Boswell, J. F., Gallagher, M. W., Sauer-Zavala, S. E., Bullis, J., Gorman, J. M., Shear, M. K., Woods, S., & Barlow, D. H. (2013). Patient characteristics and variability in adherence and competence in cognitive-behavioral therapy for panic disorder. *Journal of Consulting and Clinical Psychology, 81*(3), 443.

Boudreaux, M. J., Ozer, D. J., Oltmanns, T. F., & Wright, A. G. (2018). Development and validation of the Circumplex Scales of Interpersonal Problems. *Psychological Assessment, 30*(5), 594.

Brotman, M. A. (2004). *Therapeutic alliance and adherence in cognitive therapy for depression* [Doctoral thesis, University of Pennsylvania]. Dissertation Abstracts International, 65(3146), 6B. (UMI No. 3169565).

Brown, D. P. (2006). *Pointing out the great way: The stages of meditation in the Mahamudra tradition*. Simon and Schuster.

Budd, R., & Hughes, I. (2009). The Dodo Bird Verdict—controversial, inevitable and important: a commentary on 30 years of meta-analyses. *Clinical Psychology & Psychotherapy: An International Journal of Theory & Practice, 16*(6), 510-522.

Burlingame, G. M., MacKenzie, K., & Strauss, B. (2004). Small group treatment: Evidence for effectiveness and mechanisms of change. In M. J. Lambert (Ed.), *Handbook of psychotherapy and behavior change* (5 ed., pp. 647–696). Wileys & Sons.

Burns, D. D., & Auerbach, A. (1996). Therapeutic empathy in cognitive-behavioral therapy: does it really make a difference? In P. Salkovskis (Ed.), (pp. 135–164). Frontiers of cognitive therapy, Guilford Press.

Busse, R., Kratochwill, T. R., & Elliott, S. N. (1995). Meta-analysis for single-case consultation outcomes: Applications to research and practice. *Journal of School Psychology, 33*(4), 269-285.

Butler, S. F., & Strupp, H. H. (1986). Specific and nonspecific factors in psychotherapy: A problematic paradigm for psychotherapy research. *Psychotherapy: Theory, Research, Practice, Training, 23*(1), 30.

Carlyle, D., Green, R., Inder, M., Porter, R., Crowe, M., Mulder, R., & Frampton, C. (2020). A Randomized-Controlled Trial of Mentalization-Based Treatment Compared With Structured Case Management for Borderline Personality Disorder in a Mainstream Public Health Service. *Frontiers in Psychiatry, 11*, 1252.

Castonguay, L., Constantino, M., McAleavey, A., & Goldfried, M. (2010). The therapeutic alliance in cognitive-behavioral therapy. In J. Muran & J. Barber (Eds.), *The therapeutic alliance: An evidence-based guide to practice* (pp. 150–171). New York: Guilford Press.

Castonguay, L. G. (1993). "Common factors" and "nonspecific variables": Clarification of the two concepts and recommendations for research. *Journal of Psychotherapy Integration, 3*(3), 267.

Castonguay, L. G., Goldfried, M. R., Wiser, S., Raue, P. J., & Hayes, A. M. (1996). Predicting the effect of cognitive therapy for depression: a study of unique and common factors. *Journal of Consulting and Clinical Psychology, 64*(3), 497.

Castonguay, L. G., & Grosse, M. (2005). Change in psychotherapy: a plea for no more "nonspecific" and false dichotomies. *Clinical Psychology: Science and Practice, 12*(2), 198-201.

Chambless, D., & Crits-Christoph, P. (2006). *The treatment method.* American Psychologicsl Association.

Chambless, D. L., Crits-Christoph, P., Wampold, B. E., Norcross, J. C., Lambert, M. J., Bohart, A. C., Beutler, L. E., & Johannsen, B. E. (2006). What should be validated? In *Evidence-based practices in mental health: Debate and dialogue on the fundamental questions.* (pp. 191-256). American Psychological Association.

Chambless, D. L., & Ollendick, T. H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual review of psychology, 52*(1), 685-716.

Chapman, C. L., Baker, E. L., Porter, G., Thayer, S. D., & Burlingame, G. M. (2010). Rating group therapist interventions: The validation of the Group Psychotherapy Intervention Rating Scale. *Group Dynamics: Theory, Research, and Practice, 14*(1), 15.

Chiesa, M., Luyten, P., & Fonagy, P. (2020). Two-year follow-up and changes in reflective functioning in specialist and nonspecialist treatment models for personality disorder. *Personality Disorders: Theory, Research, and Treatment*, No Pagination Specified-No Pagination Specified.

Choi-Kain, L. W., Reich, D. B., Masland, S. R., Iliakis, E. A., & Ilagan, G. S. (2020). Longitudinal Course of Borderline Personality Disorder: What Every Clinician Needs to Know. *Current Treatment Options in Psychiatry, 7*(3), 429–445.

Cicchetti, D., & Rogosch, F. A. (2002). A developmental psychopathology perspective on adolescence. *Journal of Consulting and Clinical Psychology, 70*(1), 6.

Clarkin, J. F., & Levy, K. N. (2004). The influence of client variables on psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5 ed., pp. 194–226). Wiley.

Clarkin, J. F., Levy, K. N., Lenzenweger, M. F., & Kernberg, O. F. (2007). Evaluating three treatments for borderline personality disorder: A multiwave study. *American Journal of Psychiatry, 164*(6), 922-928.

Cloitre, M., Chase Stovall-McClough, K., Miranda, R., & Chemtob, C. M. (2004). Therapeutic alliance, negative mood regulation, and treatment outcome in child abuse-related posttraumatic stress disorder. *Journal of Consulting and Clinical Psychology, 72*(3), 411–416.

Cohen, J. N., McElhaney, R., & Jensen, D. (2018). The empirical examinability of psychodynamic psychotherapy: A reply to Hoffart and Johnson. *Clinical Psychological Science, 6*(4), 458-463.

Colli, A., Tanzilli, A., Dimaggio, G., & Lingiardi, V. (2014). Patient personality and therapist response: An empirical investigation. *American Journal of Psychiatry, 171*(1), 102-108.

Cologon, J., Schweitzer, R. D., King, R., & Nolte, T. (2017). Therapist reflective functioning, therapist attachment style and therapist effectiveness. *Administration and Policy in Mental Health and Mental Health Services Research, 44*(5), 614-625.

Cook, T. D., Campbell, D. T., & Day, A. (1979). *Quasi-experimentation: Design & analysis issues for field settings* (Vol. 351). Houghton Mifflin Boston.

Cosin, B., Freeman, C., & Freeman, N. (1971). Critical empiricism criticized: the case of Freud. *Journal for the Theory of Social Behaviour, 1*(2), 121-151.

Coupland, C. A., Hill, T., Dening, T., Morriss, R., Moore, M., & Hippisley-Cox, J. (2019). Anticholinergic drug exposure and the risk of dementia: a nested case-control study. *JAMA internal medicine, 179*(8), 1084-1093.

Cramer, V., Torgersen, S., & Kringlen, E. (2006). Personality disorders and quality of life. A population study. *Comprehensive psychiatry, 47*(3), 178-184.

Crawford, M. J., Thana, L., Farquharson, L., Palmer, L., Hancock, E., Bassett, P., Clarke, J., & Parry, G. D. (2016). Patient experience of negative effects of psychological treatment: results of a national survey. *The British Journal of Psychiatry, 208*(3), 260-265.

Cristea, I. A., Gentili, C., Cotet, C. D., Palomba, D., Barbui, C., & Cuijpers, P. (2017). Efficacy of psychotherapies for borderline personality disorder: a systematic review and meta-analysis. *Jama psychiatry, 74*(4), 319-328.

Crits-Christoph, P., Baranackie, K., Kurcias, J., Beck, A., Carroll, K., Perry, K., Luborsky, L., McLellan, A., Woody, G., & Thompson, L. (1991). Meta-analysis of therapist effects in psychotherapy outcome studies. *Psychotherapy research, 1*(2), 81-91.

Crits-Christoph, P., Chambless, D. L., & Markell, H. M. (2014). Moving evidence-based practice forward successfully: Commentary on Laska, Gurman, and Wampold. *Psychotherapy, 51*(4), 491–495.

Crits-Christoph, P., & Connolly, M. B. (1999). Alliance and technique in short-term dynamic therapy. *Clinical Psychology Review, 19*(6), 687-704.

Crits-Christoph, P., Cooper, A., & Luborsky, L. (1988). The accuracy of therapists' interpretations and the outcome of dynamic psychotherapy. *Journal of Consulting and Clinical Psychology, 56*(4), 490.

Crits-Christoph, P., Gallop, R., Temes, C. M., Woody, G., Ball, S. A., Martino, S., & Carroll, K. M. (2009). The alliance in motivational enhancement therapy and counseling as usual for substance use problems. *Journal of Consulting and Clinical Psychology, 77*(6), 1125.

Crits-Christoph, P., Gibbons, M. B. C., Crits-Christoph, K., Narducci, J., Schamberger, M., & Gallop, R. (2006a). Can therapists be trained to improve their alliances? A preliminary study of alliance-fostering psychotherapy. *Psychotherapy research, 16*(03), 268-281.

Crits-Christoph, P., Gibbons, M. B. C., & Hearon, B. (2006b). Does the alliance cause good outcome? Recommendations for future research on the alliance. *Psychotherapy: Theory, Research, Practice, Training, 43*(3), 280–285.

Crits-Christoph, P., Johnson, J., Gallop, R., Gibbons, M. B. C., Ring-Kurtz, S., Hamilton, J. L., & Tu, X. (2011). A generalizability theory analysis of group process ratings in the treatment of cocaine dependence. *Psychotherapy research, 21*(3), 252-266.

Cronbach, L. J., Rajaratnam, N., & Gleser, G. C. (1963). Theory of generalizability: A liberalization of reliability theory. *British Journal of Statistical Psychology, 16*(2), 137-163.

Csibra, G., & Gergely, G. (2006). Social learning and social cognition: The case for pedagogy. In Y. Munakata & M. H. Johnson (Eds.), *Processes of Change in Brain and Cognitive Development (Attention and Performance Vol. 21)* (pp. 249–274). Oxford University Press.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in cognitive sciences, 13*(4), 148-153.

Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences, 366*(1567), 1149-1157.

Dalio, R. (2017). *Principles: Life and work*. Simon and Schuster.

Daniell, D. (1985). Love and work: complementary aspects of personal identity. *International Journal of Social Economics, 12*(2), 48–55.

Darchuk, A., Wang, V., Weibel, D., Fende, J., Anderson, T., & Horvath, A. (2000). Manual for the Working Alliance Inventory—Observer form, 4th revision. *Unpublished manuscript.*

Davanloo, H. (1990a). The technique of unlocking the unconscious in patients suffering from functional disorders. Part I. Restructuring ego's defenses. *Davanloo H. Unlocking the unconscious. Chichester: John Wiley & Sons*, 283-306.

Davanloo, H. (1990b). *Unlocking the unconscious*. Wiley.

De Bolle, M., Johnson, J. G., & De Fruyt, F. (2010). Patient and clinician perceptions of therapeutic alliance as predictors of improvement in depression. *Psychotherapy and Psychosomatics, 79*(6), 378-385.

De Meulemeester, C., Vansteelandt, K., Luyten, P., & Lowyck, B. (2018). Mentalizing as a mechanism of change in the treatment of patients with borderline personality disorder: A parallel process growth modeling approach. *Personality Disorders: Theory, Research, and Treatment, 9*(1), 22.

Derogatis, L. (2000). The Brief Symptom Inventory-18 (BSI-18): Administration. *Scoring, and Procedures Manual (3rd ed.), Minneapolis: National Computer Systems.*

Derogatis, L. R. (2017). Symptom Checklist-90-Revised, Brief Symptom Inventory, and BSI-18. In M. E. Maruish (Ed.), *Handbook of Psychological Assessment in Primary Care Settings* (2 ed., pp. 599–629). Routledge/Taylor & Francis Group.

Derogatis, L. R., & Unger, R. (2010). Symptom checklist-90-revised. In N. C. & C. W. E. (Eds.), *The Corsini encyclopedia of psychology* (pp. 1–2). Hoboken, NJ: John Wiley & Sons.

DeRubeis, R. J., Brotman, M. A., & Gibbons, C. J. (2005). A conceptual and methodological analysis of the nonspecifics argument. *Clinical Psychology: Science and Practice, 12*(2), 174-183.

DeRubeis, R. J., & Feeley, M. (1990). Determinants of change in cognitive therapy for depression. *Cognitive therapy and research, 14*(5), 469-482.

Descartes, R. (1960). Discourse on method and meditations (LJ Lafleur, Trans.). *Indianapolis: Bobbs-Merrill.*

Dimaggio, G., Semerari, A., Carcione, A., Nicolō, G., & Procacci, M. (2007). *Psychotherapy of personality disorders: Metacognition, states of mind and interpersonal cycles*. Routledge.

Dinger, U., Strack, M., Leichsenring, F., Wilmers, F., & Schauenburg, H. (2008). Therapist effects on outcome and alliance in inpatient psychotherapy. *Journal of clinical psychology, 64*(3), 344-354.

Ditlefsen, I. T., Nissen-Lie, H., Andenæs, A., Normann-Eide,E., Johansen, M. S., Kvarstein, E.H. (2020). "Yes, there is actually hope!"– A qualitative investigation of how patients experience mentalization-based psychoeducation tailored for borderline personality disorder. *Journal of Psychotherapy Integration*.

Dolan, R. T., Arnkoff, D. B., & Glass, C. R. (1993). Client attachment style and the psychotherapist's interpersonal stance. *Psychotherapy: Theory, Research, Practice, Training, 30*(3), 408.

Duncan, B. L., Miller, S. D., Wampold, B. E., & Hubble, M. A. (2010). *The heart and soul of change: Delivering what works in therapy*. American Psychological Association.

Eatough, V., & Smith, J. A. (2008). Interpretative phenomenological analysis. In C. Willig & W. Stainton-Rogers (Eds.), *The Sage handbook of qualitative research in psychology* (pp. 179–194). Thousand Oaks: Sage.

Elkin, I. (1999). A major dilemma in psychotherapy outcome research: Disentangling therapists from therapies. *Clinical Psychology: Science and Practice, 6*(1), 10-32.

Elkin, I., Falconnier, L., Martinovich, Z., & Mahoney, C. (2006). Therapist effects in the national institute of mental health treatment of depression collaborative research program. *Psychotherapy research, 16*(02), 144-160.

Elliott, R. (2002). Hermeneutic single-case efficacy design. *Psychotherapy research, 12*(1), 1-21.

Elliott, R. (2010). Psychotherapy change process research: Realizing the promise. *Psychotherapy research, 20*(2), 123-135.

Elliott, R. (2011). Qualitative methods for studying psychotherapy change processes. In *Qualitative research methods in mental health & psychotherapy: A Guide for Students and Practitioners* (pp. 69-81). Wiley-Blackwell.

Ellison, W. D. (2020). Psychotherapy for Borderline Personality Disorder: Does the Type of Treatment Make a Difference? *Current Treatment Options in Psychiatry, 7*(3), 416-428.

Ellison, W. D., Rosenstein, L. K., Morgan, T. A., & Zimmerman, M. (2018). Community and clinical epidemiology of borderline personality disorder. *Psychiatric Clinics, 41*(4), 561-573.

Enck, P., & Zipfel, S. (2019). Placebo effects in psychotherapy: a framework. *Frontiers in Psychiatry, 10*, 456.

Endicott, J., Spitzer, R. L., Fleiss, J. L., & Cohen, J. (1976). The Global Assessment Scale: A procedure for measuring overall severity of psychiatric disturbance. *Archives of general psychiatry, 33*(6), 766-771.

Esposito, G., Passeggia, R., Cutolo, A. S., Karterud, S., & Freda, M. F. (2020). Treatment integrity and members' change in group counseling: A pilot study on counselor's mentalizing interventions. *Professional psychology: Research and practice, 51*(6), 588–597.

Eubanks-Carter, C., Muran, J. C., & Safran, J. D. (2010). Alliance ruptures and resolution. In J. C. Muran & J. P. Barber (Eds.), *The therapeutic alliance: An evidence-based guide to practice* (pp. 74–94). Guilford.

Euler, S., Stalujanis, E., Allenbach, G., Kolly, S., De Roten, Y., Despland, J.-N., & Kramer, U. (2019). Dialectical behavior therapy skills training affects defense mechanisms in borderline personality disorder: An integrative approach of mechanisms in psychotherapy. *Psychotherapy research, 29*(8), 1074-1085.

Fairbairn, W. R. D. (1954). *An object-relations theory of the personality*. Basic Books.

Falkenström, F. (2003). A Buddhist contribution to the psychoanalytic psychology of self. *The International Journal of Psychoanalysis, 84*(6), 1551-1568.

Falkenström, F. (2010). Studying mindfulness in experienced meditators: A quasi-experimental approach. *Personality and Individual Differences, 48*(3), 305-310.

Falkenström, F., Granström, F., & Holmqvist, R. (2013a). Therapeutic alliance predicts symptomatic improvement session by session. *Journal of counseling psychology, 60*(3), 317–328.

Falkenström, F., Hatcher, R. L., & Holmqvist, R. (2015). Confirmatory factor analysis of the patient version of the Working Alliance Inventory–Short Form Revised. *Assessment, 22*(5), 581-593.

Falkenström, F., & Larsson, M. H. (2017). The working alliance: From global outcome prediction to micro-analyses of within-session fluctuations. *Psychoanalytic Inquiry, 37*(3), 167-178.

Falkenström, F., Markowitz, J. C., Jonker, H., Philips, B., & Holmqvist, R. (2013b). Can psychotherapists function as their own controls? Meta-analysis of the "crossed therapist" design in comparative psychotherapy trials. *The Journal of clinical psychiatry, 74*(5), 482.

Falkenström, F., Solbakken, O. A., Möller, C., Lech, B., Sandell, R., & Holmqvist, R. (2014). Reflective functioning, affect consciousness, and mindfulness: Are these different functions? *Psychoanalytic psychology, 31*(1), 26.

Feenstra, D. J., Hutsebaut, J., Verheul, R., & Busschbach, J. J. (2011). Severity Indices of Personality Problems (SIPP–118) in adolescents: Reliability and validity. *Psychological Assessment, 23*(3), 646.

Feynman, M. (2018). *Growing Up Feynman – Michelle Feynman – 5/11/2018*. YouTube. Retrieved 28.01.2021 from https://www.youtube.com/watch?v=GnSvy3nH7l0

Feynman, R. P. (2006). *QED: The strange theory of light and matter*. Princeton University Press.

Finlay, L., & Gough, B. (2008). *Reflexivity: A practical guide for researchers in health and social sciences*. John Wiley & Sons.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. (1994). Structured clinical interview for Axis I DSM-IV disorders. *New York: Biometrics Research*.

Fischer-Kern, M., Doering, S., Taubner, S., Hörz, S., Zimmermann, J., Rentrop, M., Schuster, P., Buchheim, P., & Buchheim, A. (2015). Transference-focused psychotherapy for borderline personality disorder: Change in reflective function. *The British Journal of Psychiatry, 207*(2), 173-174.

Fitzmaurice, G., Davidian, M., Verbeke, G., & Molenberghs, G. (2008). *Longitudinal data analysis*. CRC press.

Flückiger, C., Del Re, A., Wampold, B. E., & Horvath, A. O. (2018). The alliance in adult psychotherapy: A meta-analytic synthesis. *Psychotherapy, 55*(4), 316.

Folmo, E., Karterud, S., Kongerslev, M., Kvarstein, E., & Stänicke, E. (2019). Battles of the comfort zone: modelling therapeutic strategy, alliance, and epistemic trust—a qualitative study of mentalization-based therapy for borderline personality disorder. *Journal of Contemporary Psychotherapy, 49*(3), 141-151.

Folmo, E., Stänicke, E., Johansen, M., Pedersen, G., & Kvarstein, E. (2020). Development of therapeutic alliance in Mentalization-based treatment. *Manuscript submitted for publication*.

Folmo, E. J., & Langjord, T. (2014). *Affective Activation but not Insight Predicts Sense of Self in Short-Term Dynamic Psychotherapy: A Single-case Study* [Master's thesis, University of Oslo].

Folmo, E. J., Langjord, T., Stänicke, E., & Kvarstein, E. H. (2021a). Pedagogic interventions in Mentalization-based treatment (MBT). *Manuscript in preparation*.

Folmo, E. J., Langjord, T., Stänicke, E., & Kvarstein, E. H. (2021b). Pedagogic stance in Mentalization-based treatment. *Manuscript in preparation*.

Fonagy, P. (1999). Achieving evidence-based psychotherapy practice: A psychodynamic perspective on the general acceptance of treatment manuals. *Clinical Psychology: Science and Practice, 6*(4), 442-444.

Fonagy, P. (2010). The changing shape of clinical practice: Driven by science or by pragmatics? *Psychoanalytic Psychotherapy, 24*(1), 22-43.

Fonagy, P., & Allison, E. (2014). The role of mentalizing and epistemic trust in the therapeutic relationship. *Psychotherapy, 51*(3), 372.

Fonagy, P., & Bateman, A. W. (2006). Mechanisms of change in mentalization-based treatment of BPD. *Journal of clinical psychology, 62*(4), 411-430.

Fonagy, P., & Bateman, A. W. (2007). Mentalizing and borderline personality disorder. *Journal of Mental Health, 16*(1), 83-101.

Fonagy, P., Campbell, C., Truscott, A., & Fuggle, P. (2020). Debate: Mentalising remotely– The AFNCCF's adaptations to the coronavirus crisis. *Child and Adolescent Mental Health, 25*(3), 178–179.

Fonagy, P., Gergely, G., Target, M., & Jurist, E. L. (2002). *Affect Regulation, Mentalization, and the Development of the Self*. Other Press, LLC.

Fonagy, P., Luyten, P., & Allison, E. (2015a). Epistemic petrification and the restoration of epistemic trust: A new conceptualization of borderline personality disorder and its psychosocial treatment. *Journal of personality disorders, 29*(5), 575-609.

Fonagy, P., Luyten, P., Allison, E., & Campbell, C. (2019). Mentalizing, epistemic trust and the phenomenology of psychotherapy. *Psychopathology, 52*(2), 94-103.

Fonagy, P., Luyten, P., & Bateman, A. (2015b). Translation: Mentalizing as treatment target in borderline personality disorder. *Personality Disorders: Theory, Research, and Treatment, 6*(4), 380–392.

Fonagy, P., Target, M., Steele, H., & Steele, M. (1998). *Reflective-functioning manual, version 5.0, for application to adult attachment interviews*. London: University College London. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.466.3207&rep=rep1&type=pdf

Fosha, D. E., Siegel, D. J., & Solomon, M. F. (2009). *The healing power of emotion: Affective neuroscience, development & clinical practice*. WW Norton & Company.

Fraley, R., Waller, N., & Brennan, K. (2000). The Experiences in Close Relationships-Revised (ECR-R) Questionnaire. *Downloaded from: http://internal.psychology.illinois.edu/~rcfraley/measures/ecrritems.htm*.

Frank, J. D., & Frank, J. B. (1993). *Persuasion and healing: A comparative study of psychotherapy*. JHU Press.

Frank, J. D., Nash, E. H., Stone, A. R., & Imber, S. D. (1963). Immediate and long-term symptomatic course of psychiatric outpatients. *American Journal of Psychiatry, 120*(5), 429-439.

Freud, S. (1961). The future of an illusion, civilization and its discontents, and other works. In J. Strachey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 21, pp. 5–58). Hogarth Press.

Gabbard, G. O., Horwitz, L., Frieswyk, S., Allen, J. G., Colson, D. B., Newsom, G., & Coyne, L. (1988). The effect of therapist interventions on the therapeutic alliance with borderline patients. *Journal of the American Psychoanalytic Association, 36*(3), 697-727.

Gadamer, H.-G. (1975). Hermeneutics and social science. *Cultural hermeneutics, 2*(4), 307-316.

Gadamer, H.-G., Weinsheimer, J., & Marshall, D. G. (2004). *EPZ truth and method*. Bloomsbury Publishing USA.

Garfield, S. L. (1978). Research on client variables in psychotherapy. In *Handbook of psychotherapy and behavior change* (pp. 191–232). John Wiley & Sons Inc.

Gaston, L. (1990). The concept of the alliance and its role in psychotherapy: Theoretical and empirical considerations. *Psychotherapy: Theory, Research, Practice, Training, 27*(2), 143–153.

Gaston, L., Thompson, L., Gallagher, D., Cournoyer, L.-G., & Gagnon, R. (1998). Alliance, technique, and their interactions in predicting outcome of behavioral, cognitive, and brief dynamic therapy. *Psychotherapy research, 8*(2), 190-209.

Gelso, C. J., & Carter, J. A. (1985). The relationship in counseling and psychotherapy: Components, consequences, and theoretical antecedents. *The Counseling Psychologist, 13*(2), 155-243.

Gibbons, M. C., Crits-Christoph, P., de la Cruz, C., Barber, J. P., Siqueland, L., & Gladis, M. (2003). Pretreatment expectations, interpersonal functioning, and symptoms in the prediction of the therapeutic alliance across supportive-expressive psychotherapy and cognitive therapy. *Psychotherapy research, 13*(1), 59-76.

Giesen-Bloo, J., Van Dyck, R., Spinhoven, P., Van Tilburg, W., Dirksen, C., Van Asselt, T., Kremers, I., Nadort, M., & Arntz, A. (2006). Outpatient psychotherapy for borderline personality disorder: randomized trial of schema-focused therapy vs transference-focused psychotherapy. *Archives of general psychiatry, 63*(6), 649-658.

Goldberg, S. B., Flemotomos, N., Martinez, V. R., Tanana, M. J., Kuo, P. B., Pace, B. T., Villatte, J. L., Georgiou, P. G., Van Epps, J., & Imel, Z. E. (2020). Machine learning and natural language processing in psychotherapy research: Alliance as example use case. *Journal of counseling psychology, 67*(4), 438.

Goldfried, M. R. (1980). Toward the delineation of therapeutic change principles. *American Psychologist, 35*(11), 991.

Goldstein, R. E. (2015). *Promoting reflection in psychotherapy: Patient attachment, patient and trainee therapist reflective functioning, affective word use, and symptom change*

in psychodynamic therapy sessions [Doctoral dissertation, Department of Psychology, Long Island University, Brooklyn, New York].

Goodman, G. (2013). Is mentalization a common process factor in transference-focused psychotherapy and dialectical behavior therapy sessions? *Journal of Psychotherapy Integration, 23*(2), 179.

Grawe, K. (1997). Research-informed psychotherapy. *Psychotherapy research, 7*(1), 1–19.

Greenberg, L. S. (2007). A guide to conducting a task analysis of psychotherapeutic change. *Psychotherapy research, 17*(1), 15-30.

Greenson, R. R. (1965). THE WORKING ALLIANCE AND THE TRANSFERENCE NEUROSIS. *The Psychoanalytic Quarterly, 34*, 155.

Grissom, R. J. (1996). The magical number. 7±. 2: Meta-meta-analysis of the probability of superior outcome in comparisons involving therapy, placebo, and control. *Journal of Consulting and Clinical Psychology, 64*(5), 973.

Gullestad, F. S., & Wilberg, T. (2011). Change in reflective functioning during psychotherapy—A single-case study. *Psychotherapy research, 21*(1), 97-111.

Gunderson, J., Masland, S., & Choi-Kain, L. (2018a). Good psychiatric management: a review. *Current Opinion in Psychology, 21*, 127-131.

Gunderson, J., Najavits, L., Leonhard, C., Sullivan, C., & Sabo, A. (1997). Ontogeny of the therapeutic alliance in borderline patients. *Psychotherapy research, 7*(3), 301-309.

Gunderson, J. G., Herpertz, S. C., Skodol, A. E., Torgersen, S., & Zanarini, M. C. (2018b). Borderline personality disorder. *Nature Reviews Disease Primers, 4*(1), 1-20.

Guntrip, H. (1973). *Psychoanalytic theory, therapy, and the self: A guide to the human personality in Freud, Erikson, Klein, Sullivan, Fairbairn, Hartmann, Jacobson, and Winnicott*. New York: Basic Books. (Original woik published 1971).

Gutermann, J., Schreiber, F., Matulis, S., Stangier, U., Rosner, R., & Steil, R. (2015). Therapeutic adherence and competence scales for Developmentally Adapted Cognitive Processing Therapy for adolescents with PTSD. *European Journal of Psychotraumatology, 6*(1), 26632.

Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für mathematik und physik, 38*(1), 173-198.

Habermas, J. (1986). *Knowledge and human interests*. Polity.

Hagtvet, K. A. (1997). The Function of Indicators and Errors in Construct Measures: An Application of Generalizability Theory. *Journal of Vocational Education Research, 22*(4), 247-266.

Hamilton, E., Cairns, H., & Cooper, L. (1961). *The collected dialogues of Plato*. Princeton University Press.

Harding, S. (1975). *Can theories be refuted?: Essays on the Duhem-Quine thesis* (Vol. 81). Springer Science & Business Media.

Hatcher, B. A. (1999). *Swami in Wonderland*. Oxford University Press on Demand.

Hatcher, R. L. (2010). *Alliance theory and measurement*. Guilford Press.

Hatcher, R. L., Barends, A., Hansell, J., & Gutfreund, M. J. (1995). Patients' and therapists' shared and unique views of the therapeutic alliance: An investigation using confirmatory factor analysis in a nested design. *Journal of Consulting and Clinical Psychology, 63*(4), 636–643.

Hatcher, R. L., & Barends, A. W. (2006). How a return to theory could help alliance research. *Psychotherapy: Theory, Research, Practice, Training, 43*(3), 292–299.

Hatcher, R. L., & Gillaspy, J. A. (2006). Development and validation of a revised short version of the Working Alliance Inventory. *Psychotherapy research, 16*(1), 12-25.

Haynes, S. N., & O'Brien, W. H. (2003). *Principles and practice of behavioral assessment*. Springer Science & Business Media.

Hedeker, D., & Gibbons, R. D. (1997). Application of random-effects pattern-mixture models for missing data in longitudinal studies. *Psychological methods, 2*(1), 64–78.

Hegel, G. W. F. (2010). *Georg Wilhelm Friedrich Hegel: the science of logic*. Cambridge University Press.

Hegel, G. W. F. (2018). *Hegel: The phenomenology of spirit*. Oxford University Press.

Heidegger, M. (1976). *What Is Called Thinking?* Harper Perennial.

Hembree, E. A., Rauch, S. A., & Foa, E. B. (2003). Beyond the manual: The insider's guide to prolonged exposure therapy for PTSD. *Cognitive and Behavioral Practice, 10*(1), 22-30.

Henry, W. P., Schacht, T. E., & Strupp, H. H. (1986). Structural analysis of social behavior: Application to a study of interpersonal process in differential psychotherapeutic outcome. *Journal of Consulting and Clinical Psychology, 54*(1), 27–31.

Henry, W. P., & Strupp, H. H. (1994). The therapeutic alliance as interpersonal process. In A. O. Horvath & L. S. Greenberg (Eds.), *The working alliance: Theory, research, and practice* (pp. 51–84). Wiley.

Henry, W. P., Strupp, H. H., Butler, S. F., Schacht, T. E., & Binder, J. L. (1993). Effects of training in time-limited dynamic psychotherapy: Changes in therapist behavior. *Journal of Consulting and Clinical Psychology, 61*(3), 434.

Hilliard, R. B. (1993). Single-case methodology in psychotherapy process and outcome research. *Journal of Consulting and Clinical Psychology, 61*(3), 373–380.

Hobbs, N. (1962). Sources of gain in psychotherapy. *American Psychologist, 17*(11), 741.

Hoeflin, R. K. (2004). *To Unscrew the Inscrutable*. Published by Ronald K. Hoeflin.

Hoffart, A., Borge, F.-M., Sexton, H., Clark, D. M., & Wampold, B. E. (2012). Psychotherapy for social phobia: How do alliance and cognitive process interact to produce outcome? *Psychotherapy research, 22*(1), 82-94.

Hoffart, A., & Johnson, S. U. (2017). Psychodynamic and cognitive-behavioral therapies are more different than you think: conceptualizations of mental problems and consequences for studying mechanisms of change. *Clinical Psychological Science, 5*(6), 1070-1086.

Holtforth, M. G., & Castonguay, L. G. (2005). Relationship and techniques in cognitive-behavioral therapy--A motivational approach. *Psychotherapy: Theory, Research, Practice, Training, 42*(4), 443.

Horvath, A., & Bedi, R. (2002). The alliance. In J. C. Norcross (Ed.), *Psychotherapy relationships that work* (pp. 37–69). Oxford University Press.

Horvath, A. O. (2001). The alliance. *Psychotherapy: Theory, Research, Practice, Training, 38*(4), 365–372.

Horvath, A. O. (2006). The alliance in context: Accomplishments, challenges, and future directions. *Psychotherapy: Theory, Research, Practice, Training, 43*(3), 258.

Horvath, A. O. (2018). Research on the alliance: Knowledge in search of a theory. *Psychotherapy research, 28*(4), 499-516.

Horvath, A. O., Del Re, A., Flückiger, C., & Symonds, D. (2011). Alliance in individual psychotherapy. *Psychotherapy, 48*(1), 9.

Horvath, A. O., & Greenberg, L. S. (1989). Development and validation of the Working Alliance Inventory. *Journal of counseling psychology, 36*(2), 223.

Horvath, A. O., & Luborsky, L. (1993). The role of the therapeutic alliance in psychotherapy. *Journal of Consulting and Clinical Psychology, 61*(4), 561.

Horvath, A. O., & Symonds, B. D. (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of counseling psychology, 38*(2), 139.

Huppert, J. D., Bufka, L. F., Barlow, D. H., Gorman, J. M., Shear, M. K., & Woods, S. W. (2001). Therapists, therapist variables, and cognitive-behavioral therapy outcome in a multicenter trial for panic disorder. *Journal of Consulting and Clinical Psychology, 69*(5), 747.

Hutsebaut, J., Bales, D. L., Busschbach, J. J., & Verheul, R. (2012). The implementation of mentalization-based treatment for adolescents: a case study from an organizational, team and therapist perspective. *International Journal of Mental Health Systems, 6*(1), 1-11.

Høglend, P. (2014). Exploration of the patient-therapist relationship in psychotherapy. *American Journal of Psychiatry, 171*(10), 1056-1066.

Høglend, P., Amlo, S., Marble, A., Bøgwald, K.-P., Sørbye, Ø., Sjaastad, M. C., & Heyerdahl, O. (2006). Analysis of the patient-therapist relationship in dynamic psychotherapy: An experimental study of transference interpretations. *American Journal of Psychiatry, 163*(10), 1739-1746.

Haaga, D. A., & Stiles, W. B. (2000). Randomized clinical trials in psychotherapy research: Methodology, design, and evaluation. In C. R. Snyder & R. E. Ingram (Eds.), *Handbook of psychological change* (pp. 14–39). Wiley.

Iacoviello, B. M., McCarthy, K. S., Barrett, M. S., Rynn, M., Gallop, R., & Barber, J. P. (2007). Treatment preferences affect the therapeutic alliance: implications for randomized controlled trials. *Journal of Consulting and Clinical Psychology, 75*(1), 194.

Imel, Z. E., Baer, J. S., Martino, S., Ball, S. A., & Carroll, K. M. (2011). Mutual influence in therapist competence and adherence to motivational enhancement therapy. *Drug and Alcohol Dependence, 115*(3), 229-236.

Imel, Z. E., & Wampold, B. E. (2008). The importance of treatment and the science of common factors in psychotherapy. In S. D. Brown & R. W. Lent (Eds.), *Handbook of counseling psychology* (4 ed., pp. 249–262). Hoboken, NJ: John Wiley and Sons.

Inderhaug, T. S. (2013). *Challenges in Mentalization–Based Group Therapy (MBT-G): A Video-based Explorative Case Study* [Master thesis, The Norwegian University of Science and Technology]. Trondheim.

Inderhaug, T. S., & Karterud, S. (2015). A qualitative study of a mentalization-based group for borderline patients. *Group Analysis, 48*(2), 150-163.

Iwakabe, S., & Gazzola, N. (2009). From single-case studies to practice-based knowledge: Aggregating and synthesizing case studies. *Psychotherapy research, 19*(4-5), 601-611.

Jacobson, N. S., & Truax, P. (1991). Clinical Significance: A Statistical Approach to Denning Meaningful Change in Psychotherapy Research. *Journal of Consulting and Clinical Psychologv, 59*(1), 12-19.

Jarrett, C. (2008). When therapy causes harm. *Psychologist, 21*(1), 10-12.

Johansen, S. E. (2008). *Grunnriss av en differensiell epistemologi (English translation from Norwegian in transit: Outline of Differential Epistemology)* (2 ed.). Abstrakt.

Johns, R. G., Barkham, M., Kellett, S., & Saxon, D. (2019). A systematic review of therapist effects: A critical narrative update and refinement to review. *Clinical Psychology Review, 67*, 78-93.

Jones, E. E. (1985). Psychotherapy process Q-set coding manual. *University of California, Berkeley*.

Jørgensen, C. R. (2004). Active Ingredients in Individual Psychotherapy: Searching for Common Factors. *Psychoanalytic psychology, 21*(4), 516–540.

Jørgensen, C. R. (2019). Randomized Controlled Trials. In *The Psychotherapeutic Stance* (pp. 53-70). Springer.

Jørgensen, C. R., Bøye, R., Andersen, D., Døssing Blaabjerg, A. H., Freund, C., Jordet, H., & Kjølbye, M. (2014). Eighteen months post-treatment naturalistic follow-up study of mentalization-based therapy and supportive group treatment of borderline personality disorder: Clinical outcomes and functioning. *Nordic Psychology, 66*(4), 254-273.

Jørgensen, C. R., Freund, C., Bøye, R., Jordet, H., Andersen, D., & Kjølbye, M. (2013). Outcome of mentalization-based and supportive psychotherapy in patients with borderline personality disorder: a randomized trial. *Acta Psychiatrica Scandinavica, 127*(4), 305-317.

Jørgensen, M. S., Storebø, O. J., Bo, S., Poulsen, S., Gondan, M., Beck, E., Chanen, A. M., Bateman, A., Pedersen, J., & Simonsen, E. (2020). Mentalization-based treatment in groups for adolescents with borderline personality disorder: 3-and 12-month follow-up of a randomized controlled trial. *European child and adolescent psychiatry*.

Kalleklev, J., & Karterud, S. (2018). A comparative study of a mentalization-based versus a psychodynamic group therapy session. *Group Analysis, 51*(1), 44-60.

Kant, I. (2007). Critique of pure reason (M. Weigelt, Trans.). *UK: Penguin Classics.(Original work published, 1781, 1787)*.

Kaptchuk, T. J., Kelley, J. M., Conboy, L. A., Davis, R. B., Kerr, C. E., Jacobson, E. E., Kirsch, I., Schyner, R. N., Nam, B. H., & Nguyen, L. T. (2008). Components of placebo effect: randomised controlled trial in patients with irritable bowel syndrome. *Bmj, 336*(7651), 999-1003.

Karlsson, R., & Kermott, A. (2006). Reflective-functioning during the process in brief psychotherapies. *Psychotherapy: Theory, Research, Practice, Training, 43*(1), 65.

Karterud, S. (2011). *Manual for mentaliseringsbasert psykoedukativ gruppeterapi (MBT-I)*. Gyldendal akademisk.

Karterud, S. (2012). *Manual for mentaliseringsbasert gruppeterapi (MBT-G)*. Gyldendal akademisk Oslo.

Karterud, S. (2015). *Mentalization-Based Group Therapy (MBT-G): A theoretical, clinical, and research manual*. OUP Oxford.

Karterud, S. (2017). *Personlighet*. Gyldendal akademisk.

Karterud, S. (2018). Case formulations in mentalization-based group therapy. *Research in Psychotherapy: Psychopathology, Process and Outcome, 21*(3).

Karterud, S. (2019). *Manual for gruppekurs om personlighet og personlighetsforstyrrelser*. Gyldendal Akademisk.

Karterud, S., & Bateman, A. (2010). *Manual for mentaliseringsbasert terapi (MBT) og MBT vurderingsskala. Versjon individualterapi*. Oslo: Gyldendal akademisk.

Karterud, S., Folmo, E., & Kongerslev, M. (2019). Personality and the group matrix. *Group Analysis, 52*(4), 503-519.

Karterud, S., Folmo, E., & Kongerslev, M. (2020). *Mentaliseringsbasert terapi MBT*. Gyldendal Akademisk.

Karterud, S., & Kongerslev, M. T. (2019a). Case formulations in mentalization-based treatment (MBT) for patients with borderline personality disorder. In U. Kramer (Ed.), *Case formulation for personality disorders: Tailoring psychotherapy to the individual client* (pp. 41–60). Academic Press.

Karterud, S., Pedersen, G., Engen, M., Johansen, M. S., Johansson, P. N., Schlüter, C., Urnes, Ø., Wilberg, T., & Bateman, A. W. (2013). The MBT Adherence and Competence Scale (MBT-ACS): development, structure and reliability. *Psychotherapy research, 23*(6), 705-717.

Karterud, S., Pedersen, G., Johansen, M., Wilberg, T., Davis, K., & Panksepp, J. (2016). Primary emotional traits in patients with personality disorders. *Personality and mental health, 10*(4), 261-273.

Karterud, S. W., & Kongerslev, M. T. (2019b). A temperament-attachment-mentalization-based (TAM) theory of personality and its disorders. *Frontiers in psychology, 10*, 518.

Katz, M., Hilsenroth, M. J., Gold, J. R., Moore, M., Pitman, S. R., Levy, S. R., & Owen, J. (2019). Adherence, flexibility, and outcome in psychodynamic treatment of depression. *66*(1), 94–103.

Katznelson, H. (2014). Reflective functioning: A review. *Clinical Psychology Review, 34*(2), 107-117.

Kazdin, A. (1978). *History of behavior modification: Experimental foundations of contemporary research.(xi ed.) Baltimore*. MD, US: University Park Press.

Kazdin, A., & Tuma, A. (1982). *New directions for methodology of social and behavioral sciences: Single-case research designs*. San Francisco: Jossey-Bass.

Kazdin, A. E. (1998). *Research design in clinical psychology*. Allyn & Bacon.

Kazdin, A. E. (2005). Treatment outcomes, common factors, and continued neglect of mechanisms of change. *Clinical Psychology: Science and Practice, 12*(2), 184-188.

Kazdin, A. E. (2006). Arbitrary metrics: implications for identifying evidence-based treatments. *American Psychologist, 61*(1), 42.

Kazdin, A. E. (2009). Understanding how and why psychotherapy leads to change. *Psychotherapy research, 19*(4-5), 418-428.

Kazdin, A. E. (2011). Evidence-based treatment research: Advances, limitations, and next steps. *American Psychologist, 66*(8), 685.

Kegan, R., & Lahey, L. L. (2009). *Immunity to Change: How to Overcome it and Unlock Potential in Yourself and Your Organization*. Harvard Business Press. https://books.google.no/books?id=rbhlTRx5VusC

Kermode, F., & Kermode, J. F. (1975). *Selected prose of TS Eliot*. London: Faber.

Kernberg, O. F. (2016). What is personality? *Journal of personality disorders, 30*(2), 145-156.

Kernberg, O. F., & Caligor, E. (2005). A psychoanalytic theory of personality disorders. *Major theories of personality disorder, 2*, 114-156.

Khan, A., Redding, N., & Brown, W. A. (2008). The persistence of the placebo response in antidepressant clinical trials. *Journal of psychiatric research, 42*(10), 791-796.

Kierkegaard, S. (1992). *Concluding unscientific postscript to "Philosophical Fragments"*. Princeton University Press. (1846)

Kierkegaard, S. (1995). Works of love (HV Hong & EH Hong, eds. & trans.). *Princeton, New Jersey: Princeton Univer-sity Press.(Original work published 1847)*.

Kierkegaard, S. (1998). *The point of view* (Vol. 22). Princeton University Press.

Kierkegaard, S. (2013). *Kierkegaard's Writings, II, Volume 2: The Concept of Irony, with Continual Reference to Socrates/Notes of Schelling's Berlin Lectures* (Vol. 76). Princeton University Press.

Kiesler, D. J. (1986). Interpersonal methods of diagnosis and treatment. *Psychiatry, 1*(4), 1-23.

Kiesler, D. J. (1995). Research classic:"Some myths of psychotherapy research and the search for a paradigm": Revisited. *Psychotherapy research, 5*(2), 91-101.

Killingmo, B., Varvin, S., & Strømme, H. (2014). What can we expect from trainee therapists? A study of acquisition of competence in dynamic psychotherapy. *The Scandinavian Psychoanalytic Review, 37*(1), 24-35.

Kirsch, I., Wampold, B., & Kelley, J. M. (2016). Controlling for the placebo effect in psychotherapy: Noble quest or tilting at windmills? *Psychology of Consciousness: Theory, Research, and Practice, 3*(2), 121.

Kivity, Y., Levy, K. N., Wasserman, R. H., Beeney, J. E., Meehan, K. B., & Clarkin, J. F. (2019). Conformity to prototypical therapeutic principles and its relation with change in reflective functioning in three treatments for borderline personality disorder. *Journal of Consulting and Clinical Psychology, 87*(11), 975.

Klein, D. N., Schwartz, J. E., Santiago, N. J., Vivian, D., Vocisano, C., Castonguay, L. G., Arnow, B., Blalock, J. A., Manber, R., & Markowitz, J. C. (2003). Therapeutic alliance in depression treatment: controlling for prior change and patient characteristics. *Journal of Consulting and Clinical Psychology, 71*(6), 997–1006.

Kohut, H. (1984). How does analysis cure Chicago. *IL: University of Chicago*.

Kuhn, T. S. (1970). *Criticism and the growth of knowledge: Volume 4: Proceedings of the International Colloquium in the Philosophy of Science, London, 1965* (Vol. 4). Cambridge University Press.

Kuhn, T. S. (1996). *The Structure of Scientific Revolutions. revised and augmented*. Chicago and London: University of Chicago Press.

Kvarstein, E. (2013). *Psychotherapy of personality disorder: Large variations in the severity and longitudinal course of global functioning, symptom distress, and costs of health services* [Doctoral dissertation, University of Oslo].

Kvarstein, E. H., Folmo, E., Antonsen, B. T., Normann-Eide, E., Pedersen, G., & Wilberg, T. (2020). Social cognition capacities as predictors of outcome in Mentalization-based treatment (MBT). *Frontiers in Psychiatry, 11*, 691.

Kvarstein, E. H., Pedersen, G., Folmo, E., Urnes, Ø., Johansen, M. S., Hummelen, B., Wilberg, T., & Karterud, S. (2019). Mentalization-based treatment or psychodynamic treatment programmes for patients with borderline personality disorder–the impact of clinical severity. *Psychology and Psychotherapy: Theory, Research and Practice, 92*(1), 91-111.

Kvarstein, E. H., Pedersen, G., Urnes, Ø., Hummelen, B., Wilberg, T., & Karterud, S. (2015). Changing from a traditional psychodynamic treatment programme to mentalization-based treatment for patients with borderline personality disorder–Does it make a difference? *Psychology and Psychotherapy: Theory, Research and Practice, 88*(1), 71-86.

Lakatos, I. (1970). Falsificationism and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 91–195). Cambridge University Press.

Lakatos, I., Worrall, J., & Currie, G. (1980). *The methodology of scientific research programmes: Volume 1: Philosophical papers (Vol. 1)*. Cambridge University Press.

Lambert, M. (1992). Implications of psychotherapy outcome research for eclectic and integrative psychotherapies. *Handbook of Psychotherapy Integration. New York: Basic Books*.

Lambert, M., & Ogles, B. (2004). *The efficacy and effectiveness of psychotherapy. In, MJ Lambert (Ed.) Bergin and Garfields handbook of psychotherapy and behavior change (pp. 139-193)*. New York: Wiley.

Lambert, M. J. (1989). The individual therapist's contribution to psychotherapy process and outcome. *Clinical Psychology Review, 9*(4), 469-485.

Lambert, M. J. (2013). *Bergin and Garfield's handbook of psychotherapy and behavior change*. John Wiley & Sons.

Lambert, M. J., & Ogles, B. M. (2009). Using clinical significance in psychotherapy outcome research: The need for a common procedure and validity data. *Psychotherapy research, 19*(4-5), 493-501.

Langjord, T. (2009). *"And the pool was filled with water out of sunlight": Nagarjuna og nondualitet i TS Eliots Four Quartets* [Master thesis, University of Oslo].

Langton, C. G. (1986). Studying artificial life with cellular automata. *Physica D: Nonlinear Phenomena, 22*(1-3), 120-149.

Laska, K., & Wampold, B. (2014a). Ten things to remember about common factor theory. *Psychotherapy (Chicago, Ill.), 51*(4), 519.

Laska, K. M., Gurman, A. S., & Wampold, B. E. (2014). Expanding the lens of evidence-based practice in psychotherapy: A common factors perspective. *Psychotherapy, 51*(4), 467.

Laska, K. M., & Wampold, B. E. (2014b). Ten things to remember about common factor theory. *Psychotherapy, 51*(4), 519–524.

Laurenssen, E., Luyten, P., Kikkert, M. J., Westra, D., Peen, J., Soons, M. B., van Dam, A.-M., van Broekhuyzen, A. J., Blankers, M., & van Busschbach, J. (2018). Day hospital mentalization-based treatment versus specialist treatment as usual in patients with borderline personality disorder: Randomized controlled trial. *Psychological medicine, 48*(15), 2522-2529.

Laurenssen, E. M., Eeren, H. V., Kikkert, M. J., Peen, J., Westra, D., Dekker, J. J., & Busschbach, J. J. (2016). The burden of disease in patients eligible for mentalization-based treatment (MBT): quality of life and costs. *Health and quality of life outcomes, 14*(1), 1-9.

Leichsenring, F., Leibing, E., Kruse, J., New, A. S., & Leweke, F. (2011). Borderline personality disorder. *The Lancet, 377*(9759), 74-84.

Leichsenring, F., Steinert, C., & Crits-Christoph, P. (2018). On mechanisms of change in psychodynamic therapy. *Zeitschrift für Psychosomatische Medizin und Psychotherapie, 64*(1), 16-22.

Lemma, A., Target, M., & Fonagy, P. (2011). *Brief dynamic interpersonal therapy: A clinician's guide*. Oxford University Press.

Levy, K., N, Meehan, K., B, Reynoso, J., S, Lenzenweger, M., F, Clarkin, J., F, & Kernberg, O., F. (2005). Abstracts of the 2005 Poster Session of the American Psychoanalytic Association Winter Meeting: The Relation of Reflective Funct. *Journal of the American Psychoanalytic Association, 53*(4), 1305-1308.

Levy, K. N. (2008). Psychotherapies and Lasting Change. *The American journal of psychiatry, 165*(5), 556-559.

Levy, K. N., Beeney, J. E., Wasserman, R. H., & Clarkin, J. F. (2010). Conflict begets conflict: Executive control, mental state vacillations, and the therapeutic alliance in treatment of borderline personality disorder. *Psychotherapy research, 20*(4), 413-422.

Levy, K. N., Meehan, K. B., Kelly, K. M., Reynoso, J. S., Weber, M., Clarkin, J. F., & Kernberg, O. F. (2006). Change in attachment patterns and reflective function in a randomized control trial of transference-focused psychotherapy for borderline personality disorder. *Journal of Consulting and Clinical Psychology, 74*(6), 1027.

Lilja, J. L., Lundh, L.-G., Josefsson, T., & Falkenström, F. (2013). Observing as an essential facet of mindfulness: a comparison of FFMQ patterns in meditating and non-meditating individuals. *Mindfulness, 4*(3), 203-212.

Lincoln, Y., & Guba, E. (1985). Naturalistic inquiry.,(Sage Publications: Newbury Park, CA).

Lind, M., Jørgensen, C. R., Heinskou, T., Simonsen, S., Bøye, R., & Thomsen, D. K. (2019a). Patients with borderline personality disorder show increased agency in life stories after 12 months of psychotherapy. *Psychotherapy, 56*(2), 274.

Lind, M., Thomsen, D. K., Bøye, R., Heinskou, T., Simonsen, S., & Jørgensen, C. R. (2019b). Personal and parents' life stories in patients with borderline personality disorder. *Scandinavian journal of psychology, 60*(3), 231-242.

Linden, M., & Schermuly-Haupt, M.-L. (2014). Definition, assessment and rate of psychotherapy side effects. *World Psychiatry, 13*(3), 306.

Lindgren, A., Barber, J. P., & Sandahl, C. (2008). Alliance to the Group–as–a–Whole as a Predictor of Outcome in Psychodynamic Group Therapy. *International journal of group psychotherapy, 58*(2), 163-184.

Linehan, M. (1993). *Cognitive-behavioral Treatment of Borderline Personality Disorder*. Guilford Press.

Linehan, M. (2014). *DBT? Skills training manual*. Guilford Publications.

Lingiardi, V., Filippucci, L., & Baiocco, R. (2005). Therapeutic alliance evaluation in personality disorders psychotherapy. *Psychotherapy research, 15*(1-2), 45-53.

Luborsky, L. (1976). Helping alliances in psychotherapy: The groundwork for a study of their relationship to its outcome. In J. L. Cleghhorn (Ed.), *Successful psychotherapy* (pp. 92–116). Brunner/Mazel.

Luborsky, L., & Bachrach, H. (1974). Factors influencing clinician's judgments of mental health: Eighteen experiences with the Health-Sickness Rating Scale. *Archives of general psychiatry, 31*(3), 292-299.

Luborsky, L., & Barber, J. P. (1993). Benefits of adherence to psychotherapy manuals, and where to get them. In N. E. Miller, L. Luborsky, J. P. Barber, & J. P. Docherty (Eds.), *Psychodynamic treatment research: A handbook for clinical practice* (pp. 211–226). Basic Books.

Luborsky, L., Crits-Christoph, P., Alexander, L., Margolis, M., & Cohen, M. (1983). Two helping alliance methods for predicting outcomes of psychotherapy: A counting signs vs. a global rating method. *Journal of nervous and mental disease.*

Luborsky, L., Crits-Christoph, P., McLellan, A. T., Woody, G., Piper, W., Liberman, B., Imber, S., & Pilkonis, P. (1986). DO THERAPISTS VARY MUCH IN THEIR SUCCESS?. Findings from Four Outcome Studies. *American Journal of Orthopsychiatry, 56*(4), 501-512.

Luborsky, L., Diguer, L., Seligman, D. A., Rosenthal, R., Krause, E. D., Johnson, S., Halperin, G., Bishop, M., Berman, J. S., & Schweizer, E. (1999). The researcher's own therapy allegiances: A "wild card" in comparisons of treatment efficacy. *Clinical Psychology: Science and Practice, 6*(1), 95-106.

Luborsky, L., Rosenthal, R., Diguer, L., Andrusyna, T. P., Berman, J. S., Levitt, J. T., Seligman, D. A., & Krause, E. D. (2002). The dodo bird verdict is alive and well—mostly. *Clinical Psychology: Science and Practice, 9*(1), 2-12.

Lutz, W., Leon, S., Martinovich, Z., Lyons, J., & Stiles, W. (2007). Therapist effects and therapist performance clusters in outpatient psychotherapy. *Journal of counseling psychology, 54*, 32-39.

Luyten, P., Campbell, C., Allison, E., & Fonagy, P. (2020a). The mentalizing approach to psychopathology: State of the art and future directions. *Annual Review of Clinical Psychology, 16*, 297–325.

Luyten, P., Campbell, C., & Fonagy, P. (2020b). Borderline personality disorder, complex trauma, and problems with self and identity: A social-communicative approach. *Journal of personality, 88*(1), 88–105.

Magnusson, E., & Marecek, J. (2015). *Doing interview-based qualitative research: A learner's guide*. Cambridge University Press.

Malda-Castillo, J., Browne, C., & Perez-Algorta, G. (2019). Mentalization-based treatment and its evidence-base status: A systematic literature review. *Psychology and Psychotherapy: Theory, Research and Practice, 92*(4), 465-498.

Mallinckrodt, B. (2000). Attachment, social competencies, social support, and interpersonal process in psychotherapy. *Psychotherapy research, 10*(3), 239-266.

Martin, D. J., Garske, J. P., & Davis, M. K. (2000). Relation of the therapeutic alliance with outcome and other variables: a meta-analytic review. *Journal of Consulting and Clinical Psychology, 68*(3), 438.

Martinez, R. G. (2020). *A meta-analysis investigating the correlation between treatment integrity and youth client outcomes* [Doctoral dissertation, Virginia Commonwealth University].

Marziali, E., Munroe-Blum, H., & McCleary, L. (1999). The effects of the therapeutic alliance on the outcomes of individual and group psychotherapy with borderline personality disorder. *Psychotherapy research, 9*(4), 424-436.

Masterson, J. F. (1978). The borderline adult: therapeutic alliance and transference. *The American journal of psychiatry, 135*(4), 437–441.

Masterson, J. F. (1988). *The search for the real self: Unmasking the personality disorders of our age.* Taylor & Francis.

Mayne, T. J., & Ambrose, T. K. (1999). Research review on anger in psychotherapy. *Journal of clinical psychology, 55*(3), 353-363.

McCullough, L., Larsen, A. E., Schanche, E., Andrews, S., Kuhn, N., Hurley, C., & Wolf, J. (2003). *Achievement of Therapeutic Objectives Scale: ATOS scale.* http://www.affectphobia.org/videoatos.html

McGlinchey, J. B., & Dobson, K. S. (2003). Treatment integrity concerns in cognitive therapy for depression. *Journal of Cognitive Psychotherapy, 17*(4), 299-318.

McKay, M., Wood, J. C., & Brantley, J. (2019). *The dialectical behavior therapy skills workbook: Practical DBT exercises for learning mindfulness, interpersonal effectiveness, emotion regulation, and distress tolerance.* New Harbinger Publications.

McLeod, J. (2011). *Qualitative research in counselling and psychotherapy.* Sage.

McMain, S. F., Links, P. S., Gnam, W. H., Guimond, T., Cardish, R. J., Korman, L., & Streiner, D. L. (2009). A randomized trial of dialectical behavior therapy versus general psychiatric management for borderline personality disorder. *American Journal of Psychiatry, 166*(12), 1365-1374.

Mehlum, L., Tørmoen, A. J., Ramberg, M., Haga, E., Diep, L. M., Laberg, S., Larsson, B. S., Stanley, B. H., Miller, A. L., & Sund, A. M. (2014). Dialectical behavior therapy for adolescents with repeated suicidal and self-harming behavior: a randomized trial.

Journal of the American Academy of Child & Adolescent Psychiatry, 53(10), 1082-1091.

Messer, S., & Wolitzky, D. (2010). The therapeutic alliance: A psychodynamic perspective. In J. C. Muran & J. P. Barber (Eds.), *The therapeutic alliance: An evidence-based approach to practice and training* (pp. 97–122). Guilford Press.

Messer, S. B., & Wampold, B. E. (2002). Let's face facts: Common factors are more potent than specific therapy ingredients. *Clinical Psychology: Science and Practice, 9*(1), 21-25.

Miller, S. J., & Binder, J. L. (2002). The effects of manual-based training on treatment fidelity and outcome: A review of the literature on adult individual psychotherapy. *Psychotherapy: Theory, Research, Practice, Training, 39*(2), 184–198.

Millon , T. (2004). The antisocial personality. In T. Millon (Ed.), *Personality disorders in modern life* (pp. 150–186). Wiley & Sons.

Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological review, 102*(2), 246.

Mohr, D. C. (1995). Negative outcome in psychotherapy: A critical review. *Clinical Psychology: Science and Practice, 2*(1), 1-27.

Morken, K. T. E., Binder, P.-E., Arefjord, N. M., & Karterud, S. W. (2019). Mentalization-Based Treatment From the Patients' Perspective–What Ingredients Do They Emphasize? *Frontiers in psychology, 10*(1327).

Motesharrei, S., Rivas, J., & Kalnay, E. (2014). Human and nature dynamics (HANDY): Modeling inequality and use of resources in the collapse or sustainability of societies. *Ecological Economics, 101*, 90-102.

Munder, T., Wilmers, F., Leonhart, R., Linster, H. W., & Barth, J. (2010). Working Alliance Inventory-Short Revised (WAI-SR): psychometric properties in outpatients and inpatients. *Clinical Psychology & Psychotherapy: An International Journal of Theory & Practice, 17*(3), 231-239.

Mundt, J. C., Marks, I. M., Shear, M. K., & Greist, J. M. (2002). The Work and Social Adjustment Scale: a simple measure of impairment in functioning. *The British Journal of Psychiatry, 180*(5), 461-464.

Muran, J. C., Safran, J. D., & Eubanks-Carter, C. (2010). Developing therapist abilities to negotiate alliance ruptures. In J. C. Muran & J. P. Barber (Eds.), *The therapeutic alliance: An evidence-based guide to practice* (pp. 320–340). Guilford Press.

Muran, J. C., Safran, J. D., Gorman, B. S., Samstag, L. W., Eubanks-Carter, C., & Winston, A. (2009). The relationship of early alliance ruptures and their resolution to process and outcome in three time-limited psychotherapies for personality disorders. *Psychotherapy: Theory, Research, Practice, Training, 46*(2), 233.

Möller, C., Karlgren, L., Sandell, A., Falkenström, F., & Philips, B. (2017). Mentalization-based therapy adherence and competence stimulates in-session mentalization in psychotherapy for borderline personality disorder with co-morbid substance dependence. *Psychotherapy research, 27*(6), 749-765.

Nash, J. (1956). The imbedding problem for Riemannian manifolds. *Annals of mathematics, 63*(1), 20–63.

Nissen-Lie, H. A., Monsen, J. T., & Rønnestad, M. H. (2010). Therapist predictors of early patient-rated working alliance: A multilevel approach. *Psychotherapy research, 20*(6), 627-646.

Norcross, J., & Goldfried, M. (2005). *Oxford series in clinical psychology. Handbook of psychotherapy integration*. Oxford University Press.

Norcross, J. C., Beutler, L. E., & Levant, R. F. (2006). *Evidence-based practices in mental health: Debate and dialogue on the fundamental questions*. American Psychological Association.

Norcross, J. C., & Wampold, B. E. (2011). What works for whom: Tailoring psychotherapy to the person. *Journal of clinical psychology, 67*(2), 127-132.

Norusis, M. (2008). *SPSS 16.0 advanced statistical procedures companion*. Prentice Hall Press.

Ogles, B. M., Anderson, T., & Lunnen, K. M. (1999). The contribution of models and techniques to therapeutic efficacy: Contradictions between professional trends and clinical research. In M. A. Hubble, B. L. Duncan, & S. D. Miller (Eds.), *The heart and soul of change: What works in therapy* (pp. 201–225). American Psychological Association.

Ogrodniczuk, J. S., Piper, W. E., Joyce, A. S., & McCallum, M. (1999). Transference interpretations in short-term dynamic psychotherapy. *The journal of nervous and mental disease, 187*(9), 571-578.

Orlinsky, D., & Howard, K. (1980). Gender and psychotherapeutic outcome. In A. M. Brodsky & R. T. Hare-Mustin (Eds.), *Women and psychotherapy* (pp. 3–34). Guilford.

Orlinsky, D. E., Grawe, K., & Parks, B. K. (1994). Process and outcome in psychotherapy: noch einmal. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4 ed., pp. 270–378). Wiley.

Orlinsky, D. E., Ronnestad, M. H., & Willutzki, U. (2004). Fifty years of psychotherapy process-outcome research: Continuity and change. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5 ed., pp. 307–389). Wiley.

Oud, M., Arntz, A., Hermens, M. L., Verhoef, R., & Kendall, T. (2018). Specialized psychotherapies for adults with borderline personality disorder: a systematic review and meta-analysis. *Australian & New Zealand Journal of Psychiatry, 52*(10), 949-961.

Owen, J., & Hilsenroth, M. J. (2011). Interaction between alliance and technique in predicting patient outcome during psychodynamic psychotherapy. *The journal of nervous and mental disease, 199*(6), 384-389.

Owen, J., & Hilsenroth, M. J. (2014). Treatment adherence: The importance of therapist flexibility in relation to therapy outcomes. *Journal of counseling psychology, 61*(2), 280.

Panksepp, J. (2011). The basic emotional circuits of mammalian brains: do animals have affective lives? *Neuroscience & Biobehavioral Reviews, 35*(9), 1791-1804.

Panksepp, J., & Biven, L. (2012). *The archaeology of mind: neuroevolutionary origins of human emotions (Norton series on interpersonal neurobiology)*. WW Norton & Company.

Passer, M. W., & Smith, R. E. (2004). *Psychology: The science of mind and behavior*. McGraw-Hill.

Pedersen, G. (2008). *Psychological assessment in clinical settings. Evaluation and clinical utility of psychometric measures for the treatment of patients with personality disorders.* [Doctoral dissertation, University of Oslo. Norway].

Pedersen, G., Arnevik, E. A., Hummelen, B., Walderhaug, E., & Wilberg, T. (2017a). Psychometric properties of the severity indices of personality problems (SIPP) in two samples. *European Journal of Psychological Assessment, 35*(5), 698–711.

Pedersen, G., Hagtvet, K., & Karterud, S. (2011). Interpersonal problems: self-therapist agreement and therapist consensus. *Journal of clinical psychology, 67*(3), 308-317.

Pedersen, G., Hagtvet, K. A., & Karterud, S. (2007). Generalizability studies of the Global Assessment of Functioning–Split version. *Comprehensive psychiatry, 48*(1), 88-94.

Pedersen, G., & Karterud, S. (2004). Is SCL-90R helpful for the clinician in assessing DSM-IV symptom disorders? *Acta Psychiatrica Scandinavica, 110*(3), 215-224.

Pedersen, G., Kvarstein, E., & Wilberg, T. (2017b). The Work and Social Adjustment Scale: Psychometric properties and validity among males and females, and outpatients with and without personality disorders. *Personality and mental health, 11*(4), 215-228.

Pedersen, G., Urnes, Ø., Hummelen, B., Wilberg, T., & Kvarstein, E. (2018). Revised manual for the global assessment of functioning scale. *European Psychiatry, 51*, 16-19.

Pedersen, G. A. (2002). Norsk revidert versjon av Inventory of Interpersonal Problems—Circumplex (IIP-C). *Tidsskrift for Norsk Psykologforening, 39*(1), 25–34.

Penrose, R. (1989). The emperor's new mind Oxford University Press. *New York*.

Penrose, R. (2006). *The road to reality: A complete guide to the laws of the universe*. Random house.

Perepletchikova, F. (2007). *Treatment integrity in treatment outcome research (2000–2004): Analysis of the studies and examination of the associated factors*. Yale University.

Perepletchikova, F. (2009). Treatment integrity and differential treatment effects. *Clinical Psychology: Science and Practice, 16*(3), 379-382.

Perepletchikova, F., Treat, T. A., & Kazdin, A. E. (2007). Treatment integrity in psychotherapy research: analysis of the studies and examination of the associated factors. *Journal of Consulting and Clinical Psychology, 75*(6), 829.

Persons, J. B. (1989). *Cognitive therapy in practice: A case formulation approach*. Norton New York.

Pesale, F. P., Hilsenroth, M. J., & Owen, J. J. (2012). Patient early session experience and treatment outcome. *Psychotherapy research, 22*(4), 417-425.

Philips, B., Wennberg, P., Konradsson, P., & Franck, J. (2018). Mentalization-based treatment for concurrent borderline personality disorder and substance use disorder: a randomized controlled feasibility study. *European addiction research, 24*(1), 1-8.

Picascia, R. (2019). *Defending the Authority of Scripture: Testimony as a Source of Knowledge in Classical Indian Philosophy of Religion* [Doctoral dissertation, Harvard University, Graduate School of Arts & Sciences].

Piper, W. E., Azim, H. F., Joyce, A. S., & McCallum, M. (1991). Transference interpretations, therapeutic alliance, and outcome in short-term individual psychotherapy. *Archives of general psychiatry, 48*(10), 946-953.

Piper, W. E., & Joyce, A. S. (2001). Psychosocial treatment outcome. In W. J. Livesley (Ed.), *Handbook of personality disorders: Theory, research, and treatment* (pp. 323–343). Guilford Press.

Piper, W. E., Ogrodniczuk, J. S., Joyce, A. S., Weideman, R., & Rosie, J. S. (2007). Group composition and group therapy for complicated grief. *Journal of Consulting and Clinical Psychology, 75*(1), 116–125.

Poincaré, H. (1905). Science and Hypothesis', reprinted in H. Poincare (1913). In H. Poincaré (Ed.), *The Foundations of Science*. The Science Press.

Popper, K. (1980). *The logic of scientific discovery*. Hutchinson.

Popper, K. (2005). *The logic of scientific discovery*. Routledge.

Porter, M. E., & Teisberg, E. O. (2006). *Redefining health care: creating value-based competition on results*. Harvard business press.

Ram-Prasad, C. (2013). *Indian philosophy and the consequences of knowledge: Themes in ethics, metaphysics and soteriology*. Ashgate Publishing, Ltd.

Reddy, M. (2012). Psychotherapy-insights from bhagavad gita. *Indian journal of psychological medicine, 34*(1), 100.

Ricoeur, P. (1970). *Freud and Philosophy: an essay on interpretation*. Yale University Press.

Ricoeur, P. (1996). *The continental philosophy reader*. Psychology Press. (Original work published 1983 by Routledge , London)

Rizq, R., & Target, M. (2010). "We had a constant battle". The role of attachment status in counselling psychologists' experiences of personal therapy: Some results from a mixed-methods study. *Counselling Psychology Quarterly, 23*(4), 343-369.

Rogers, C. (2012). *Client Centered Therapy (New Ed)*. Hachette UK.

Rosenwald, G. C. (1988). A theory of multiple-case research. *Journal of personality, 56*(1), 239-264.

Rossouw, T. I., & Fonagy, P. (2012). Mentalization-based treatment for self-harm in adolescents: a randomized controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry, 51*(12), 1304-1313. e1303.

Roth, A., & Fonagy, P. (2006). *What works for whom?: a critical review of psychotherapy research*. Guilford Press.

Rousmaniere, T. (2016). *Deliberate practice for psychotherapists: A guide to improving clinical effectiveness*. Taylor & Francis.

Rowling, J. K. (2000). *Harry Potter and the Goblet of Fire*. Bloomsbury.

Rubel, J. A., Zilcha-Mano, S., Giesemann, J., Prinz, J., & Lutz, W. (2020). Predicting personalized process-outcome associations in psychotherapy using machine learning approaches—A demonstration. *Psychotherapy research, 30*(3), 300-309.

Rudden, M., Milrod, B., Target, M., Ackerman, S., & Graf, E. (2006). Reflective functioning in panic disorder patients: A pilot study. *Journal of the American Psychoanalytic Association, 54*(4), 1339-1343.

Rønnestad, M. H. (2016). Is expertise in psychotherapy a useful construct. *Psychotherapy Bulletin, 51*(1), 11-13.

Rønnestad, M. H., & Ladany, N. (2006). The impact of psychotherapy training: Introduction to the special section. *Psychotherapy research, 16*(03), 261-267.

Rønnestad, M. H., & Skovholt, T. M. (2003). The journey of the counselor and therapist: Research findings and perspectives on professional development. *Journal of career development, 30*(1), 5-44.

Rønnestad, M. H., von der Lippe, A., Axelsen, E., & Greenberg, L. S. (2006). An intensive process-outcome study of the interpersonal aspects of psychotherapy. *Oslo, Norway: Department of psychology, Univeristory of Oslo*.

Råbu, M., Halvorsen, M. S., & Haavind, H. (2011). Early relationship struggles: A case study of alliance formation and reparation. *Counselling and Psychotherapy Research, 11*(1), 23-33.

Sadler, L. S., Slade, A., & Mayes, L. C. (2006). Minding the baby: a mentalization-based parenting program. In J. G. Allen & P. Fonagy (Eds.), *Handbook of mentalization-based treatment* (pp. 271–288). John Wiley & Sons.

Safran, J. D. (2003). *Psychoanalysis and Buddhism: An unfolding dialogue*. Simon and Schuster.

Safran, J. D., & Muran, J. C. (2000). Resolving therapeutic alliance ruptures: Diversity and integration. *Journal of clinical psychology, 56*(2), 233-243.

Safran, J. D., & Wallner, L. K. (1991). The relative predictive validity of two therapeutic alliance measures in cognitive therapy. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3*(2), 188.

Schanche, E., Vøllestad, J., Binder, P. E., Osnes, B., Visted, E., Svendsen, J. L., & Sørensen, L. (2020). Can clinical psychology students benefit from brief and intensive mindfulness training? *Counselling and Psychotherapy Research, 20*(2), 311-324.

Schwandt, T. A. (2000). Three epistemological stances for qualitative inquiry: Interpretivism, hermeneutics, and social constructionism. In N. Denzin & Y. Lincoln (Eds.), *The Landscape of Qualitative Research: Theories and Issues* (2 ed., pp. 189–214). Thousand Oaks, CA: Sage.

Schwartz, D. (2011). An Aristotelian view of knowledge for knowledge management. In *Encyclopedia of Knowledge Management, Second Edition* (pp. 39-48). IGI Global.

Sehgal, M. (2019). *Rooh-e-Rumi: Seeking God is Seeking Love*. Notion Press. https://books.google.no/books?id=2nOiDwAAQBAJ

Sharp, C., Shohet, C., Givon, D., Penner, F., Marais, L., & Fonagy, P. (2020). Learning to mentalize: A mediational approach for caregivers and therapists. *Clinical Psychology: Science and Practice, 27*(3), Article e12334.

Sharp, C., Wright, A. G., Fowler, J. C., Frueh, B. C., Allen, J. G., Oldham, J., & Clark, L. A. (2015). The structure of personality pathology: Both general ('g') and specific ('s') factors? *Journal of abnormal psychology, 124*(2), 387.

Shavelson, R. J., & Webb, N. M. (1991). *Generalizability theory: A primer* (Vol. 1). Sage.

Shavelson, R. J., Webb, N. M., & Rowley, G. L. (1989). Generalizability theory. *American Psychologist, 44*(6), 922.

Shedler, J. (2010). The efficacy of psychodynamic psychotherapy. *American Psychologist, 65*(2), 98.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., & Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry, 59*(Suppl 20), 22–33.

Shida, T. (2011). Hypothesis-Generating Logic in Udayana's Rational Theology. *Journal of Indian Philosophy, 39*(4-5), 503-520.

Shrout, P. E. (1998). Measurement reliability and agreement in psychiatry. *Statistical methods in medical research, 7*(3), 301-317.

Silbersweig, D., Clarkin, J. F., Goldstein, M., Kernberg, O. F., Tuescher, O., Levy, K. N., Brendel, G., Pan, H., Beutel, M., & Pavony, M. T. (2007). Failure of frontolimbic inhibitory function in the context of negative emotion in borderline personality disorder. *American Journal of Psychiatry, 164*(12), 1832-1841.

Simonsen, S., Juul, S., Kongerslev, M., Bo, S., Folmo, E., & Karterud, S. (2019). The mentalization-based therapy adherence and quality scale (MBT-AQS): Reliability in a clinical setting. *Nordic Psychology, 71*(2), 104-115.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford university press.

Smith, J. A. (2011). Evaluating the contribution of interpretative phenomenological analysis: A reply to the commentaries and further development of criteria. *Health psychology review, 5*(1), 55-61.

Smith, J. E., Flower, P., & Larkin, M. (2009). *Interpretative Phenomenological Analysis: Theory, Method and Research*. Sage.

Smith, M. L., Glass, G. V., & Miller, T. I. (1980). *The benefits of psychotherapy*. Johns Hopkins University Press.

Soeteman, D. I., Verheul, R., & Busschbach, J. J. (2008). The burden of disease in personality disorders: diagnosis-specific quality of life. *Journal of personality disorders, 22*(3), 259-268.

Solbakken, O. A., & Abbass, A. (2014). Implementation of an intensive short-term dynamic treatment program for patients with treatment-resistant disorders in residential care. *BMC psychiatry, 14*(1), 1-11.

Solbakken, O. A., Hansen, R. S., Havik, O. E., & Monsen, J. T. (2011). Assessment of affect integration: Validation of the affect consciousness construct. *Journal of personality assessment, 93*(3), 257-265.

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language, 25*(4), 359-393.

Spinhoven, P., Giesen-Bloo, J., van Dyck, R., Kooiman, K., & Arntz, A. (2007). The therapeutic alliance in schema-focused therapy and transference-focused psychotherapy for borderline personality disorder. *Journal of Consulting and Clinical Psychology, 75*(1), 104.

Steinberg, L., & Avenevoli, S. (2000). The role of context in the development of psychopathology: A conceptual framework and some speculative propositions. *Child development, 71*(1), 66-74.

Stiles, W. B. (2005). Case studies. In J. C. Norcross, L. E. Beutler, & R. F. Levant (Eds.), *Evidence-based practices in mental health: Debate and dialogue on the fundamental questions* (pp. 57–64). American Psychological Association.

Stiles, W. B. (2007). Theory-building case studies of counselling and psychotherapy. *Counselling and Psychotherapy Research, 7*(2), 122-127.

Stiles, W. B. (2009). Logical operations in theory-building case studies. *Pragmatic case studies in psychotherapy, 5*(3), 9-22.

Stiles, W. B., & Goldsmith, J. Z. (2010). The alliance over time. In J. C. Muran & J. P. Barber (Eds.), *The therapeutic alliance: An evidence-based guide to practice* (pp. 44–62). The Guilford Press.

Stiles, W. B., Honos-Webb, L., & Surko, M. (1998). Responsiveness in psychotherapy. *Clinical Psychology: Science and Practice, 5*(4), 439-458.

Stiles, W. B., Shapiro, D. A., & Elliott, R. (1986). Are all psychotherapies equivalent? *American Psychologist, 41*(2), 165–180.

Storebø, O. J., Stoffers-Winterling, J. M., Völlm, B. A., Kongerslev, M. T., Mattivi, J. T., Jørgensen, M. S., Faltinsen, E., Todorovac, A., Sales, C. P., & Callesen, H. E. (2020). Psychological therapies for people with borderline personality disorder (Protocol). *Cochrane Database of Systematic Reviews, 2, CD012955*.

Strupp, H. H. (1984). *Psychotherapy in a new key: A guide to time-limited dynamic psychotherapy*. Basic Books (AZ).

Strupp, H. H. (1986). The nonspecific hypothesis of therapeutic effectiveness: A current assessment. *American Journal of Orthopsychiatry, 56*(4), 513-520.

Svartberg, M., Stiles, T. C., & Seltzer, M. H. (2005). Randomized, controlled trial of the effectiveness of short-term dynamic psychotherapy and cognitive therapy for cluster C personality disorders. *Focus, 161*(3), 810-416.

Saarni, S. I., Suvisaari, J., Sintonen, H., Koskinen, S., Härkänen, T., & Lönnqvist, J. (2007). The health-related quality-of-life impact of chronic conditions varied with age in

general population. *Journal of Clinical Epidemiology, 60*(12), 1288. e1281-1288. e1211.

Tang, Y.-Y., & Posner, M. I. (2009). Attention training and attention state training. *Trends in cognitive sciences, 13*(5), 222-227.

Tasca, G. A., & Lampard, A. M. (2012). Reciprocal influence of alliance to the group and outcome in day treatment for eating disorders. *Journal of counseling psychology, 59*(4), 507–517.

Taubner, S., Kessler, H., Buchheim, A., Kächele, H., & Staun, L. (2011). The role of mentalization in the psychoanalytic treatment of chronic depression. *Psychiatry: Interpersonal & Biological Processes, 74*(1), 49-57.

Taubner, S., & Volkert, J. (2019). Evidence-Based Psychodynamic Therapies for the Treatment of Patients With Borderline Personality Disorder. *Clinical Psychology in Europe, 1*(2), 1-20.

Taubner, S., White, L. O., Zimmermann, J., Fonagy, P., & Nolte, T. (2013). Attachment-related mentalization moderates the relationship between psychopathic traits and proactive aggression in adolescence. *Journal of abnormal child psychology, 41*(6), 929-938.

Tmej, A., Fischer-Kern, M., Doering, S., Alexopoulos, J., & Buchheim, A. (2018). Changes in attachment representation in psychotherapy: is reflective functioning the crucial factor? *Zeitschrift für Psychosomatische Medizin und Psychotherapie, 64*(3), 222-236.

Tmej, A., Fischer-Kern, M., Doering, S., Hörz-Sagstetter, S., Rentrop, M., & Buchheim, A. (2020). Borderline patients before and after one year of transference-focused psychotherapy (TFP): A detailed analysis of change of attachment representations. *Psychoanalytic psychology, 38*(1), 12–21.

Tomkins, S. S. (1992). *Affect, imagery, consciousness, Vol. 4: Cognition: Duplication and transformation of information*. Springer Publishing Co.

Torgersen, S., Kringlen, E., & Cramer, V. (2001). The prevalence of personality disorders in a community sample. *Archives of general psychiatry, 58*(6), 590-596.

Trowell, J., Davids, Z., Miles, G., Shmueli, A., & Paton, A. (2008). Developing healthy mental health professionals: what can we learn from trainees? *Infant Observation, 11*(3), 333–343.

Tschacher, W., Junghan, U. M., & Pfammatter, M. (2014). Towards a taxonomy of common factors in psychotherapy—results of an expert survey. *Clinical psychology & psychotherapy, 21*(1), 82-96.

Turing, A. M. (1936). On computable numbers, with an application to the Entscheidungsproblem. *J. of Math, 58*(345-363), 5.

Ulvenes, P. G., Berggraf, L., Hoffart, A., Levy, R. A., Ablon, J. S., McCullough, L., & Wampold, B. E. (2012a). Can two psychotherapy process measures be dependably rated simultaneously? A generalizability study. *Journal of counseling psychology, 59*(4), 638.

Ulvenes, P. G., Berggraf, L., Hoffart, A., Stiles, T. C., Svartberg, M., McCullough, L., & Wampold, B. E. (2012b). Different processes for different therapies: Therapist actions, therapeutic bond, and outcome. *Psychotherapy, 49*(3), 291–302.

Valéry, P. (2015). *Collected Works of Paul Valery, Volume 14: Analects*. Princeton University Press.

Van Asselt, A., Dirksen, C., Arntz, A., Giesen-Bloo, J., & Severens, J. (2009). The EQ-5D: A useful quality of life measure in borderline personality disorder? *European Psychiatry, 24*(2), 79-85.

Venkoba, R. A., & Parvathi, D. (1974). The Bhagvad Gita treats body and mind. *Ind J Hist Med, 19*, 35-44.

Verheul, R., Andrea, H., Berghout, C. C., Dolan, C., Busschbach, J. J., van der Kroft, P. J., Bateman, A. W., & Fonagy, P. (2008). Severity Indices of Personality Problems (SIPP-118): Development, factor structure, reliability, and validity. *Psychological Assessment, 20*(1), 23.

Vermote, R., Lowyck, B., Luyten, P., Vertommen, H., Corveleyn, J., Verhaest, Y., Stroobants, R., Vandeneede, B., Vansteelandt, K., & Peuskens, J. (2010). Process and outcome in psychodynamic hospitalization-based treatment for patients with a personality disorder. *The journal of nervous and mental disease, 198*(2), 110-115.

Vogt, K. S., & Norman, P. (2019). Is mentalization-based therapy effective in treating the symptoms of borderline personality disorder? A systematic review. *Psychology and Psychotherapy: Theory, Research and Practice, 92*(4), 441-464.

Volkert, J., Hauschild, S., & Taubner, S. (2019). Mentalization-based treatment for personality disorders: Efficacy, effectiveness, and new developments. *Current psychiatry reports, 21*(4), 25.

von Consbruch, K., Clark, D. M., & Stangier, U. (2012). Assessing therapeutic competence in cognitive therapy for social phobia: psychometric properties of the cognitive therapy competence scale for social phobia (CTCS-SP). *Behavioural and Cognitive Psychotherapy, 40*(2), 149-161.

Wachtel, P. L., Kruk, J. C., & McKinney, M. K. (2005). Cyclical psychodynamics and integrative relational psychotherapy. In J. C. Nocross & M. R. Goldfried (Eds.), *Oxford series in clinical psychology: Handbook of Psychotherapy Integration* (2 ed., pp. 72–195). Oxford University Press.

Waltz, J., Addis, M. E., Koerner, K., & Jacobson, N. S. (1993). Testing the integrity of a psychotherapy protocol: assessment of adherence and competence. *Journal of Consulting and Clinical Psychology, 61*(4), 620.

Wampold, B. E. (2010). The research evidence for the common factors models: A historically situated perspective. In L. B. Duncan, S. D. Miller, B. E. Wampold, & M. A. Hubble (Eds.), *The heart and soul of change: Delivering what works in therapy* (pp. 49–81). American Psychological Association.

Wampold, B. E. (2019). *The basics of psychotherapy: An introduction to theory and practice.* American Psychological Association.

Wampold, B. E., & Bolt, D. M. (2006). Therapist effects: Clever ways to make them (and everything else) disappear. *Psychotherapy research, 16*(02), 184-187.

Wampold, B. E., & Brown, G. S. J. (2005). Estimating variability in outcomes attributable to therapists: a naturalistic study of outcomes in managed care. *Journal of Consulting and Clinical Psychology, 73*(5), 914–923.

Wampold, B. E., & Budge, S. L. (2012). The 2011 Leona Tyler Award Address: The relationship—And its relationship to the common and specific factors of psychotherapy. *The Counseling Psychologist, 40*(4), 601-623.

Wampold, B. E., Flückiger, C., Del Re, A., Yulish, N. E., Frost, N. D., Pace, B. T., Goldberg, S. B., Miller, S. D., Baardseth, T. P., & Laska, K. M. (2017). In pursuit of truth: A critical examination of meta-analyses of cognitive behavior therapy. *Psychotherapy research, 27*(1), 14-32.

Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work*. Routledge.

Wampold, B. E., Imel, Z. E., Bhati, K. S., & Johnson-Jennings, M. D. (2007a). Insight as a Common Factor. In L. G. Castonguay & C. Hill (Eds.), *Insight in psychotherapy* (pp. 119–139)). American Psychological Association.

Wampold, B. E., Imel, Z. E., & Minami, T. (2007b). The placebo effect:"Relatively large" and "robust" enough to survive another assault. *Journal of clinical psychology, 63*(4), 401-403.

Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H.-n. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empiricially," all must have prizes.". *Psychological bulletin, 122*(3), 203.

Wasserman, R. H., Levy, K. N., & Loken, E. (2009). Generalizability theory in psychotherapy research: The impact of multiple sources of variance on the dependability of psychotherapy process ratings. *Psychotherapy research, 19*(4-5), 397-408.

Watts, A. (1999). *The Philosophies of Asia: The Edited Transcripts*. C.E. Tuttle. https://books.google.no/books?id=fCacGAAACAAJ

Watts, A. (2004). Out of Your Mind: Essential Listening from the Alan Watts Audio Archives. *Sounds True, Inc. Unabridged edition.*

Webb, C. A., DeRubeis, R. J., & Barber, J. P. (2010). Therapist adherence/competence and treatment outcome: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 78*(2), 200.

Weck, F., Hilling, C., Schermelleh-Engel, K., Rudari, V., & Stangier, U. (2011). Reliability of adherence and competence assessment in cognitive behavioral therapy: Influence of clinical experience. *The journal of nervous and mental disease, 199*(4), 276-279.

Westen, D. (1995). A clinical-empirical model of personality: Life after the Mischelian ice age and the NEO-lithic era. *Journal of personality, 63*(3), 495-524.

Westen, D., & Morrison, K. (2001). A multidimensional meta-analysis of treatments for depression, panic, and generalized anxiety disorder: an empirical examination of the status of empirically supported therapies. *Journal of Consulting and Clinical Psychology, 69*(6), 875.

Wilber, K. (2000a). A brief history of everything (Rev. ed.). *Boston: Shambhala.*

Wilber, K. (2000b). *The Collected Works of Ken Wilber Vol 8: The Marriage of Sense and Soul One Taste*. Shambhala.

Wilber, K. (2001a). *Sex, ecology, spirituality: The spirit of evolution*. Shambhala Publications.

Wilber, K. (2001b). *A theory of everything: An integral vision for business, politics, science and spirituality*. Shambhala publications.

Wilkinson, R., & Pickett, K. (2011). *The spirit level: Why greater equality makes societies stronger*. Bloomsbury Publishing USA.

Winnicott, D. W. (1956). *Primary maternal preoccupation*. Tavistock.

Wittgenstein, L. (1975). *Wittgenstein's Lectures on the Foundations of Mathematics, Cambridge, 1939: From the Notes of R.G. Bosanquet, Norman Malcolm, Rush Rhees, and Yorick Smythies*. University of Chicago Press.

Wittgenstein, L. (2001). *Tractatus Logico-Philosophicus [1921], trans. Pears, D. and McGuinness, B.* Routledge, 2001.

Wittgenstein, L. (2009). *Philosophical investigations*. John Wiley & Sons.

Wittgenstein, L., Klagge, J. C., & Nordmann, A. (1993). *Philosophical Occasions, 1912-1951*. Hackett Publishing.

Wolfe, B. E., & Goldfried, M. R. (1988). Research on psychotherapy integration: Recommendations and conclusions from an NIMH workshop. *Journal of Consulting and Clinical Psychology, 56*(3), 448.

Yalom, I., & Leszcz, M. (2005). The selection of clients. In I. D. Yalom & M. Leszcz (Eds.), *Theory and practice of group psychotherapy* (5 ed., pp. 231–259). Basic Books.

Yeomans, F. E., Clarkin, J. F., & Kernberg, O. F. (2002). *A primer of transference-focused psychotherapy for the borderline patient*. Jason Aronson.

Yeomans, F. E., Gutfreund, J., Selzer, M. A., Clarkin, J. F., HULL, J. W., & SMITH, T. E. (1994). Factors related to drop-outs by borderline patients: Treatment contract and therapeutic alliance. *The Journal of psychotherapy practice and research, 3*(1), 16.

Young, J. E. (1999). *Cognitive therapy for personality disorders: A schema-focused approach*. Professional Resource Press/Professional Resource Exchange.

Young, J. E., Klosko, J. S., & Weishaar, M. E. (2006). *Schema therapy: A practitioner's guide*. Guilford Press.

Zanarini, M. C. (2003). Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD): a continuous measure of DSM-IV borderline psychopathology. *Journal of personality disorders, 17*(3), 233-242.

Zanarini, M. C., Conkey, L. C., Temes, C. M., & Fitzmaurice, G. M. (2018). Randomized, controlled trial of web-based psychoeducation for women with borderline personality disorder. *The Journal of clinical psychiatry, 79*(3).

Zanarini, M. C., & Frankenburg, F. R. (2008). A preliminary, randomized trial of psychoeducation for women with borderline personality disorder. *Journal of personality disorders, 22*(3), 284-290.

Zanarini, M. C., Frankenburg, F. R., Hennen, J., & Silk, K. R. (2004). Mental health service utilization by borderline personality disorder patients and Axis II comparison subjects followed prospectively for 6 years. *The Journal of clinical psychiatry, 65*(1), 28–36.

Zanarini, M. C., Frankenburg, F. R., Reich, D. B., Silk, K. R., Hudson, J. I., & McSweeney, L. B. (2007). The subsyndromal phenomenology of borderline personality disorder: a 10-year follow-up study. *American Journal of Psychiatry, 164*(6), 929-935.

Zanarini, M. C., Jacoby, R. J., Frankenburg, F. R., Reich, D. B., & Fitzmaurice, G. (2009). The 10-year course of social security disability income reported by patients with borderline personality disorder and axis II comparison subjects. *Journal of personality disorders, 23*(4), 346-356.

Zuroff, D. C., Kelly, A. C., Leybman, M. J., Blatt, S. J., & Wampold, B. E. (2010). Between-therapist and within-therapist differences in the quality of the therapeutic relationship: Effects on maladjustment and self-critical perfectionism. *Journal of clinical psychology, 66*(7), 681-697.

Personality and Social Psychology

The design of the MBT-G adherence and quality scale

ESPEN J. FOLMO,[1] SIGMUND W. KARTERUD,[2] KJETIL BREMER,[1] KRISTOFFER L. WALTHER,[3]
ELFRIDA H. KVARSTEIN[3,4] and GEIR A. F. PEDERSEN[3,5]

[1]Norwegian National Advisory Unit on Personality Psychiatry, Oslo University Hospital, Oslo, Norway
[2]The Norwegian Institute for Mentalizing, Oslo, Norway
[3]Department of Personality Psychiatry, Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway
[4]Institute of Clinical Medicine, University of Oslo, Norway
[5]NORMENT, KG Jebsen Center for Psychosis Research, Institute of Clinical Medicine, University of Oslo, Norway

Few group psychotherapy studies focus on therapists' interventions, and instruments that can measure group psychotherapy treatment fidelity are scarce. The aim of the present study was to evaluate the reliability of the Mentalization-based Group Therapy Adherence and Quality Scale (MBT-G-AQS), which is a 19-item scale developed to measure adherence and quality in mentalization-based group therapy (MBT-G). Eight MBT groups and eight psychodynamic groups (a total of 16 videotaped therapy sessions) were rated independently by five raters. All groups were long-term, outpatient psychotherapy groups with 1.5 hours weekly sessions. Data were analysed by a Generalizability Study (G-study and D-study). The generalizability models included analyses of reliability for different numbers of raters. The global (overall) ratings for adherence and quality showed high to excellent reliability for all numbers of raters (the reliability by use of five raters was 0.97 for adherence and 0.96 for quality). The mean reliability for all 19 items for a single rater was 0.57 (item range 0.26–0.86) for adherence, and 0.62 (item range 0.26–0.83) for quality. The reliability for two raters obtained mean absolute G-coefficients on 0.71 (item range 0.41–0.92 for the different items) for adherence and 0.76 (item range 0.42–0.91) for quality. With all five raters the mean absolute G-coefficient for adherence was 0.86 (item range 0.63–0.97) and 0.88 for quality (item range 0.64–0.96). The study demonstrates high reliability of ratings of MBT-G-AQS. In models differentiating between different numbers of raters, reliability was particularly high when including several raters, but was also acceptable for two raters. For practical purposes, the MBT-G-AQS can be used for training, supervision and psychotherapy research.

Key words: Mentalization, Mentalization-Based Treatment, group therapy, Borderline Personality Disorder, reliability, generalizability theory.

Espen J. Folmo, Department for Personality Psychiatry, Oslo University Hospital, Ullevaal, PO Box 4956 Nydalen, 0424 Oslo, Norway.
E-mail: espfol@ous-hf.no.

INTRODUCTION

Over the past decades a number of evidence-based treatment approaches for Borderline Personality Disorder (BPD) have been developed (Stoffers, Völlm, Rücker, Timmer, Huband & Lieb, 2012). One of these, mentalization-based treatment (MBT), has been found efficient in several randomised controlled trials (Bateman & Fonagy, 2001; 2009; Rossouw & Fonagy, 2012), and favorable results have been replicated in naturalistic comparisons outside the United Kingdom (Bales, Timman, Andrea, Busschbach, Verheul & Kamphuis, 2015; Kvarstein, Pedersen, Urnes, Hummelen, Wilberg & Karterud, 2015).

MBT is an intensive, combined treatment approach that includes both individual and group therapy. The four structural pillars integrated within MBT are: (1) psychoeducation; (2) an individual dynamic MBT case formulation; (3) individual mentalization-based psychotherapy (MBT-I); and (4) mentalization-based group therapy (MBT-G; Karterud, 2015). MBT thus requires a collaborative team of therapists, and the importance of regular video-based therapy supervision for MBT teams is clearly emphasized (Bateman & Fonagy, 2016).

An adherence and competence scale for MBT-I (MBT-I-ACS) has previously been developed based on a Norwegian version of the MBT manual (MBT-I; Karterud & Bateman, 2010) and the reliability of the scale was found highly satisfactory (Karterud,

Pedersen, Engen et al., 2013). The MBT-I-ACS has provided the possibility for documentation of model fidelity in studies of treatment outcomes (Kvarstein et al., 2015), and has also recently been used in a study relating outcomes to therapists' MBT interventions (Möller, Karlgren, Sandell, Falkenström & Philips, 2016).

Measures for treatment integrity are crucial when investigating whether the alleged "potion" is what is actually being delivered (Perepletchikova, Treat & Kazdin, 2007). Treatment integrity consists of two elements: (1) treatment adherence, i.e., "the extent to which a therapist used interventions and approaches prescribed by the treatment manual and avoided the use of interventions and procedures proscribed by the manual" (Waltz, Addis, Koerner & Jacobson, 1993, p. 620); and (2) the therapist's competence (quality), i.e., "the level of skill shown by the therapist in delivering the treatment" (Waltz et al., 1993, p. 620). By skill, we refer to the extent in which the therapist conducting the interventions took the relevant aspects of the therapeutic context into account and responded to these contextual variables appropriately. According to this definition, competence presupposes adherence, but adherence does not necessarily imply competence (McGlinchey & Dobson, 2003). The strong element of improvisation within dynamic psychotherapy implies that a certain competence is necessary to adhere to the ethos of the treatment. Nevertheless, such adherence can be performed with varying degrees of sophistication (timing, in-depth exploration,

integration, attunement, etc.). For the above reasons, we prefer the label 'quality' instead of competence.

Recently, both practical guidelines and manuals have been developed specifically for MBT-G (Bateman & Fonagy, 2016; Karterud, 2012, 2015). The MBT-G manual (Karterud, 2015) contains a 19-item adherence and quality scale for MBT-G (MBT-G-AQS; see Appendix).

There is a paucity of research on therapists' adherence and competence in group therapy. A review of the status of group therapy research by Burlingame, MacKenzie and Strauss (2004) issued a call for the development of group therapist intervention measures as a next step in the group treatment literature.

Documentation of treatment integrity requires manualized treatments and is essential when claiming effectiveness of specific psychotherapies (Perepletchikova *et al.*, 2007). Wampold and Imel (2015, p. 233) highlight this by stating "It is now virtually required that clinical trials of psychotherapy assess and report adherence and competence." A main challenge, present in all dynamic group therapies, is the dialectical balance between "structuring" (e.g., item 2 "Regulating group phases"; see Appendix) interventions, explorations of current mental events and overall attunement to the dynamic process (Yalom & Leszcz, 2005). The MBT-G-AQS addresses this concern through nine group-specific items and 10 further items essentially common to MBT-I-ACS.

The primary aim of the present study was to investigate the reliability of the newly developed adherence and quality scale for MBT-G. Our research questions were: (1) Can trained MBT-G raters obtain adequate interrater reliability on (a) the full MBT-G-AQS, particularly the overall ratings, and (b) adherence and quality of the nine group-specific items within MBT-G-AQS? (2) What is the minimum number of MBT-G-AQS raters required to achieve adequate reliability?

MATERIAL AND METHODS

The study is based on video-taped recordings from regular treatment groups from the same clinical unit, Department for Personality Psychiatry (DPP), Oslo University Hospital. To maximize variance, groups belonging to different time periods (2006 and 2015) were chosen. All 16 session were rated with the MBT-G-AQS.

The group therapies and group members

In the first period (2006) DPP offered a psychodynamic, group-based treatment program. In the second period (2015) MBT was the principal treatment mode. The psychodynamic group therapy (PDG) was unmanualized, followed modified group analytic principles, and was influenced by object relations theory and self-psychology (Arnevik, Wilberg, Urnes, Johansen, Monsen & Karterud, 2009). The MBT followed manual requirements as previously described (Kvarstein *et al.*, 2015).

All groups were conducted by two therapists and all group sessions lasted 1.5 hours. All groups were slow open, admitting new members whenever a place was vacant. Hence, the video material (both PDG and MBT) demonstrated patients who had attended groups for various lengths of time (range 2–36 months). Both programs combined individual and group therapy (Arnevik *et al.*, 2009; Kvarstein *et al.*, 2015).

Overall, approximately 85% of the group participants were female, age 20–30 years. The MBT groups primarily recruited BPD patients, while the PDG groups included a broader range of personality disorders (Arnevik *et al.*, 2009; Kvarstein, 2015).

Group therapists

Fourteen group therapists from the same treatment unit (57% females) participated in the study. To minimize variance due to therapists' general competence we included two therapists who performed both PDG and MBT-G. Twelve were experienced clinicians and qualified group analysts. By profession there were five psychiatrists, one psychiatric resident, two clinical psychologists, one social worker, one psychology student, one physiotherapist and three psychiatric nurses. In 2015, all therapists, except the psychiatric resident, had also received MBT training.

Scale for MBT-G

The MBT-G-AQS is a 19-item scale developed for measuring therapist adherence and quality in MBT-G. See Table 1 and Appendix for the 19 items. The manual (Karterud, 2015) contains detailed description of the development of the scale.

Video-taped group sessions

The study includes a total of 16 video-taped group therapy sessions. Eight video-tapes show PDG group sessions from 2006 and eight show MBT-G sessions from 2015. Recordings were selected by convenience sampling, i.e., aiming to minimize the variance of general therapist competence in the two time periods, 2006 and 2015. Therapist pairs in MBT and PDG were matched with respect to formal level of education. This resulted in four groups being chosen from the 2006 material. Two consecutive sessions were then selected randomly within the specified 2006 group.

The total video material from 2006 included approximately 80 sessions for each of the four PDG groups. From this pool, two consecutive sessions with the same therapist pair were randomly selected for each PDG group. In 2015, therapist-pairs from four MBT groups provided videotaped recordings of two consecutive group sessions. Two consecutive sessions were preferred in order to minimize therapists' variance over time.

MBT-G-AQS raters

Five independent clinical research collaborators rated the available video material by MBT- G- AQS (no raters were among the rated therapists). These five raters were all trained MBT therapists and familiar with MBT-I rating procedures. Prior to the current study, four of the five raters had assessed at least 30 (range 30–91) sessions with the MBT-I-ACS as part of their work for the Norwegian MBT Quality Lab. Eight hours theoretical and practical training in the MBT-G-AQS preceded the current reliability study. The pre-assessment training included rating and discussion of two verbatim transcripts of MBT groups. Four of the raters were psychologists, and one a psychiatrist (author of the MBT manual).

Table 1. *Item descriptives, G-study results, and D-study results for different measurement designs*

	Adherence / frequency								Quality								
	Grand mean[a]		G-study		D-study				Grand mean[a]		G-study		D-study				
			Coefficients		Two raters		One rater				Coefficients		Two raters		One rater		
Item	Mean	SD	Rel.[b]	Abs.[c]	Rel.[b]	Abs.[c]	Rel.[b]	Abs.[c]	Mean	SD	Rel.[b]	Abs.[c]	Rel.[b]	Abs.[c]	Rel.[b]	Abs.[c]
1. Boundaries	6.18	3.34	0.89	0.86	0.77	0.71	0.63	0.54	4.13	1.00	0.90	0.89	0.79	0.76	0.65	0.61
2. Phases	3.98	3.60	0.93	0.90	0.85	0.78	0.73	0.64	3.34	1.85	0.96	0.96	0.90	0.90	0.82	0.82
3. Turntaking	5.48	4.61	0.95	0.95	0.88	0.87	0.78	0.78	3.35	1.86	0.95	0.95	0.89	0.88	0.80	0.79
4. External events	5.38	3.44	0.94	0.92	0.86	0.83	0.75	0.71	3.38	1.44	0.93	0.92	0.85	0.83	0.74	0.71
5. Events in group	3.31	2.78	0.85	0.85	0.69	0.69	0.53	0.53	3.09	1.42	0.90	0.87	0.78	0.73	0.64	0.57
6. Care for group	Not rated								4.36	1.23	0.91	0.91	0.80	0.80	0.67	0.66
7. Authority	Not rated								4.33	1.42	0.93	0.93	0.85	0.84	0.74	0.73
8. Group norms	2.36	2.64	0.83	0.83	0.67	0.67	0.50	0.50	2.81	1.88	0.78	0.78	0.59	0.58	0.42	0.41
9. Cooperation	1.53	1.93	0.86	0.84	0.70	0.68	0.54	0.51	2.10	1.71	0.96	0.95	0.91	0.89	0.84	0.80
10. Warmth	Not rated								4.50	1.15	0.92	0.91	0.81	0.81	0.68	0.68
11. Exploration	16.03	6.79	0.87	0.80	0.73	0.61	0.58	0.44	3.98	1.32	0.86	0.86	0.71	0.71	0.56	0.55
12. Unwarranted beliefs	2.48	2.64	0.88	0.88	0.75	0.74	0.60	0.59	3.00	1.35	0.84	0.82	0.68	0.64	0.51	0.47
13. Emotional arousal	Not rated								3.68	1.29	0.93	0.93	0.84	0.84	0.72	0.72
14. Acknowledging	1.64	1.84	0.77	0.77	0.58	0.57	0.41	0.40	2.69	1.56	0.86	0.85	0.70	0.69	0.54	0.53
15. Pretend mode	Not rated								2.23	1.58	0.67	0.64	0.44	0.42	0.28	0.26
16. Psychic equivalence	1.35	1.70	0.63	0.63	0.41	0.41	0.26	0.26	2.64	1.77	0.86	0.86	0.71	0.70	0.55	0.54
17. Affect focus	14.85	7.03	0.91	0.85	0.80	0.70	0.66	0.54	4.28	1.71	0.93	0.93	0.84	0.84	0.72	0.72
18. Stop and rewind	0.65	1.08	0.88	0.88	0.74	0.74	0.59	0.59	1.80	1.81	0.85	0.84	0.69	0.67	0.53	0.50
19. Relationship	5.00	5.29	0.94	0.93	0.86	0.84	0.75	0.73	3.30	1.63	0.88	0.88	0.75	0.74	0.60	0.59
Overall rating	3.76	1.76	0.97	0.97	0.92	0.92	0.86	0.86	3.80	1.67	0.96	0.96	0.91	0.91	0.83	0.83

Notes: [a]Grand mean and standard deviations of scores across therapists and sessions. [b]Generalizability coefficient (For relative decisions). [c]Dependability coefficient (For absolute decisions).

MBT-G-AQS rating procedures

The five raters rated all MBT-G-AQS items for all 16 sessions. Ratings were performed independently, but in the same room. After having fulfilled their ratings of each session and delivered their scoring sheets to the project coordinator, the raters met and discussed agreements and disagreements, a procedure also described in other research studies (Gutermann, Schreiber, Matulis, Stangier, Rosner & Steil, 2015; von Consbruch, Clark& Stangier, 2012; Weck, Weigel, Richtberg & Stangier, 2010). Ratings were not changed after this comparison. Ratings were not blind: the raters knew most of the therapists, and were therefore not blind to treatment modality.

Ratings of adherence and competence

A therapist intervention may receive an MBT-G-AQS rating or not. A single intervention may receive more than one rating. Non-MBT interventions may sound like: "When does school start this autumn?" or "I believe the group is paralyzed for the moment" or "when did he tell you that?" Adherence on the item level is assessed by counting the frequency. Five of the items ("care for the group and its members," "managing authority," "engagement, interest and warmth," "regulating emotional arousal" and "handling pretend mode") are not assessed for adherence/frequency, as these interventions can be performed by indirect means. However, they are rated for quality. The adherence ratings equal the total number of counted interventions.

For the assessment of quality, all items are rated on a 1–7 Likert scale. The manual contains rating procedures as well as

descriptions of what counts as low versus high quality. All items are displayed in the Appendix and described by their competence level of 4 ("good enough"). If the therapists fail to deliver clearly indicated interventions, the item can be rated low on quality (e.g., 2) even where there are no occurrences. Finally, the rater decides on the overall quality score, based on a global understanding of the session.

Data analysis

In the current research design two therapy sessions from each of eight therapist-couples were videotaped. This makes a total of 16 therapy sessions, and all five raters rated all 16 sessions. In the framework of G-theory (Shavelson & Webb, 1991), this implies a two facet partially nested "(s:t) x r" design, where sessions (s) are nested within therapists (t), and raters (r) are crossed over sessions within therapists. The design is partially nested because the effect of session (s) is both nested (within t) and crossed (over r). With respect to generalizations beyond this particular study, therapists, sessions and raters are considered randomly selected from the whole 'universe' of admissible therapists, sessions and raters. The object of measurement is therapist behavior, and the measurement design is balanced as all therapists are rated by the same number of raters. The two facets of observation give two differentiation variance components, the individual variance between therapists (t) and the systematic variance between sessions for each therapist (st). This makes three sources of instrumentation variance (error) that directly effects the reliability of the observed scores. These are; (1) the rater effect (r) indicating the consistency of how much 'behavior' the
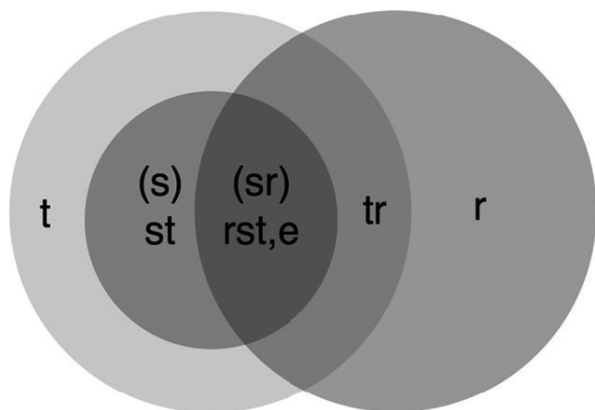
Fig. 1. Venn diagram of the variance components in the (s:t) x r design. The components are: The individual variance between therapists (t), the systematic variance between sessions for each therapist (st), the unique rater–therapist–session interaction plus other unknown error variance (rst, e), the interaction between raters and therapists (tr), and the rater effect (r). [Colour figure can be viewed at wileyonlinelibrary.com]

raters see, averaged over therapists and sessions; (2) the interaction between raters and therapists (tr), indicating the raters' different rank ordering of the therapists; and lastly (3) the unique rater–therapist–session interaction plus other unknown error variance (rst, + e) (see Fig. 1). Within this design, sessions (s) cannot be separated from therapist (t) and neither can the session–rater interaction (sr) be separated from the rater–session–therapist interaction.

Based on the sample data, the relative impact of different sources of variation is estimated by a G-study (Shavelson, Webb & Rowley, 1989), from which generalizability coefficients are computed. The G-coefficient (ρ^2) indexes the proportion of total variability in scores that is due to "universe scores" ($\rho^2 = \frac{\sigma^2(\tau)}{\sigma^2(\sigma^2(\tau)+\sigma^2(\delta)\tau)}$), where $\sigma^2(\tau)$ is the variance of the true score, and $\sigma^2(\delta)$ is the variance of the various error components. A low G-coefficient is due to a significant amount of error in measurement or to minimal variation across individuals, the measurement procedure, and the universe of generalization (Hagtvet, 1997). A G-coefficient below 0.4, is "Poor"; when it is between 0.4 and 0.59 it is "Fair"; between 0.6 and 0.74, is "Good," while a value above 0.75 is considered "Excellent" (Cicchetti, 1994).

Based on the obtained G-study components, the generalizability framework offers a subsequent study called D-study, or optimization study. Through the D-study it is possible to estimate how many conditions of each facet is necessary to obtain adequate generalizability, that is, how many raters are needed.

The intended use of the MBT-G-AQS concerns decisions of whether subjects are below or above some specific level of adherence or quality. Consequently, the most relevant reliability estimate is *absolute decisions* (i.e., absolute G-coefficients; see Karterud *et al.*, 2013 for a detailed discussion). The current G- and D-studies have been processed through the EduG program (Cardinet, Johnson & Pini, 2010; Swiss Society for Research in Education Working Group, 2010).

Ethics

After patients received a description of the study, they provided written, informed consent, as did the involved therapists. The

PDG recordings were part of the UPP project, and were approved by the Data Inspectorate and the Regional Ethics Committee in Norway. The privacy ombudsman at Oslo University Hospital approved the MBT-G part of the study.

RESULTS

Table 1 presents the reliability for all five raters and estimated D-study coefficients for two and one raters. The mean reliability on item level for adherence was 0.86 (range 0.63–0.97) and 0.88 for quality (range 0.64–0.96) for five raters. For two raters it was 0.71 (range 0.41–0.92) for adherence and 0.76 (range 0.42–0.91) for quality, which, with some exceptions, are in the acceptable to high range. The mean reliability for one rater was 0.57 (item range 0.26–0.86) for adherence, and 0.62 (item range 0.26–0.83) for quality, which ranges from poor to acceptable estimates.

The reliability for overall ratings of adherence (0.97) and quality (0.96) were both excellent. The overall ratings for adherence and quality showed high to excellent reliability for all numbers of raters. The overall ratings are also the most "immune" items for a decreasing number of raters (see Table 2). Deleting the least reliable rater from the overall ratings would only slightly increase the reliability for these two items (+0.01). This signals that the overall rating was robust also when the number of raters decreased, and that the raters agreed strongly on the overall evaluation of a MBT-G session. Table 2 demonstrates which items are most affected by a decreasing or increasing number of raters.

There were only minor differences between the reliability coefficients for absolute and relative decisions (relative and absolute G-coefficients); i.e., raters agreed as much on exact scores as on the ranking of the interventions/sessions. Therefore, all results presented are based on the absolute G-coefficients.

The nine group specific items (item 1–9) displayed very high reliability for both adherence and quality. The four items least

Table 2. *Quality rating (G-coefficient) sorted by increasing difference (5R-1R) between five raters (5R) and one rater (1R)*

Item name	5R	1R	Difference (5R-1R)
Overall rating	0.96	0.83	0.13
02. Phases	0.96	0.82	0.14
09. Cooperation	0.95	0.8	0.15
03. Turntaking	0.95	0.79	0.16
07. Authority	0.93	0.73	0.2
04. External events	0.92	0.71	0.21
13. Emotional arousal	0.93	0.72	0.21
17. Affect focus	0.93	0.72	0.21
10. Warmth	0.91	0.68	0.23
06. Care for group	0.91	0.66	0.25
01. Boundaries	0.89	0.61	0.28
19. Relationship	0.88	0.59	0.29
05. Events in group	0.87	0.57	0.3
11. Exploration	0.86	0.55	0.31
14. Acknowledging	0.85	0.53	0.32
16. Psychic equivalence	0.86	0.54	0.32
18. Stop and rewind	0.84	0.5	0.34
12. Unwarranted beliefs	0.82	0.47	0.35
08. Group norms	0.78	0.41	0.37
15. Pretend mode	0.64	0.26	0.38

affected by the number of raters decreasing, and with the highest reliability on quality, were also group specific items: "Regulating group phases," "Cooperation with cotherapist," "Initiating and fulfilling turntaking" and "Managing authority." The three group specific items "Regulating group phases," "Engaging group members in mentalizing external events" and "Initiating and fulfilling turntaking" showed very high reliability for adherence (> 0.9). "Initiating and fulfilling turntaking" was also the only item where all five raters displayed a reliability above 0.9 on adherence.

For some items the reliability would increase slightly if one of the raters was omitted in the study. These findings indicate that some of the "disagreement" on specific items was due to one rater having a different view than the others. However, there was no indication of any systemic impact on the reliability for specific raters, i.e., different raters struggled with different items.

Table 1 reveals that items 16, 14, 11 and 8 proved difficult to rate for adherence (lowest reliability). We also observe that the quality ratings for items 15, 8, 12, 18, 14, 11 and 16 were more challenging than the other items to agree on. These items had lower reliability and were also more affected by a decreasing number of raters. However, the reliability of item 16, 15, 14 and 18 is very good considering their low variance.

The two items that displayed the lowest reliability across all number of raters were "Psychic equivalence" and "Pretend mode." "Psychic equivalence" had the lowest reliability for adherence, and "Pretend mode" had the lowest reliability for quality.

From a psychometric perspective, it is ideal with some variation between therapists (T), and within therapists from session to session (T:S). Further, it is favorable that the residual variance (RS:T), raters' ranking variation (TR), and disagreement between raters (R) is as low as possible. Overall rating for quality may serve as an example of a favorable result. The residual variance for the overall quality score was very low (17%). There was complete agreement (0% variance) among the raters on how much of the intervention was observed, and the ranking of therapists. Therapists varied a lot with respect to overall quality (62% variance), but less so from session to session (21%).

From Table 1 we see that item 16 "Handling psychic equivalence" had a high residual variance (40%). There was little systematic variance between therapists regarding the intervention (11%), and from session to session (15%). There was substantial variance in the raters' ranking order (34%), but no variance in how much of the behavior (the specific intervention) the raters observed.

Item 11 ("Exploration, curiosity and not-knowing stance") had a reliability coefficient of 0.80, which is high, but low compared to the rest, especially considering high variance and frequency (mean frequency = 16). Table 3 disentangles why this particular item proved difficult to rate. Item 11 had a moderate residual variance (29% variance), which implies that the item is relatively well defined. However, there was considerable disagreement among raters on how much of this intervention they observed (24% variance), although they did not deviate much in their ranking order of the therapists (3% variance). Different opinions on what counts as item 11 interventions may have large consequences for reliability if therapist variation (between therapists and between sessions) is low. In this case, all therapists used this item frequently, as variance between therapists was very low (7%), but they varied much from session to session (37%).

Table 3 displays a relation between low reliability and residual variance. The seven items with lowest reliability had a mean residual variance of 40.5, while the seven items with highest reliability had a mean residual variance of 27. The quality ratings displayed a similar, but slightly stronger, pattern. The reason for this connection is that high residual variance signals weak references for the raters as to how to rate these items. When the residual variance for an item is high, it may indicate that therapists do not know when and how to apply it, e.g., due to poor operationalization. Hence, the item is difficult to recognize for raters.

As half of the sessions were psychodynamic groups, half of the rated therapists were not trained in the items assessed, that is, they intervened in more unfocused ways. This may explain some of the residual variance for several items: The seven items with a quality rating below 3 had a mean residual variance on adherence of 44%, while the seven items with a quality rating above 3 had a

Table 3. *Sources of variation for adherence ratings for five raters (5R): percentages of total variation. Items sorted from low to high reliability (G-coefficients; "Abs G")*

	T: between therapist variation	R: variation in how much raters observe	S:T: therapist variation across sessions	TR: variation in raters ranking of therapists	RS:T: residual (including error) variance	Abs G
16. Psychic equivalence	11.3	0	14.5	34	40.2	0.63
14. Acknowledging	23.7	2.4	16	5.2	52.6	0.77
11. Exploration	6.9	24.2	36.9	2.7	29.3	0.8
8. Group norms	7.8	0	42.1	2.6	47.6	0.83
9. Cooperation	18.8	5.2	32.7	0	43.3	0.84
17. Affect focus	20.2	19.1	33.5	1.1	26.1	0.85
5. Events in the group	16	0	37	2.5	44.5	0.85
1. Boundaries	33.1	13.5	21.4	0	32.1	0.86
18. Stop and rewind	14.7	0.6	44.2	2.5	38	0.88
12. Unwarranted beliefs	25.5	2.3	33.4	0	38.8	0.88
2. Phases	54.2	13	9.6	1.4	21.8	0.9
4. External events	50.6	5.5	20.5	5.4	17.9	0.92
19. Relationship	0	2.4	73	0	24.5	0.93
3. Turntaking	65.8	0.4	11.8	5.9	16.1	0.95

mean residual variance on adherence of 24%. The same pattern was found in the quality ratings (40/25).

DISCUSSION

This is the first study to report psychometric properties for the MBT-G-AQS. It is also the first study of a scale for measuring therapists' interventions in group therapy since 2005. The results demonstrate that the MBT-G adherence and quality scale is a reliable instrument. This scale can be applied to document treatment integrity, and underpin the evidence-base for MBT.

The overall/global ratings for adherence and quality showed high to excellent reliability across all numbers of raters. The instrument can thus be used with only one rater for research purposes where the question of overall treatment fidelity needs to be documented, and where a detailed focus on the other items are of subordinate interest. This finding also supports that the MBT-G-AQS can be reliably applied to determine if a session qualifies as "good enough" MBT-G.

At item level, the reliabilities varied substantially. This is a common finding among rating scales (Barber, Liese & Abrams, 2003). With one rater some items had a satisfactory reliability, while others had low to very low reliability. With two raters, reliabilities ranged from fair to excellent.

As process studies based on a large number of raters are very expensive and difficult to achieve (Perepletchikova, Hilt, Chereji & Kazdin, 2009), acceptable reliability for the entire scale with just one rater is important for practical implementation of the scale. Due to more extensive training, calibration, and experience of the raters in the current study, it was expected to reveal higher reliability, particularly with one and two raters, than what was obtained in the MBT-I-ACS study by Karterud *et al.* (2013). Current results confirmed this expectation. However, the reliability for one rater was still below acceptable range for several of the items. This indicates a need for further calibration and training as well as more explicit definitions of the phenomena to be assessed.

One of the benefits of performing a G-study is that it allows for identifying items that individual raters view differently than others. The finding that different raters struggled with different items means that it is important for raters to calibrate (discuss) their ratings on a regular basis.

This is particularly true for more complex (abstract) items that display low frequency, which means that raters receive less practical training in rating them. For example, "Acknowledging good mentalizing" (item 14) and "Handling psychic equivalence" (item 16) both had low frequencies, and also proved more difficult to rate for adherence than other items (with low frequency). The results indicate a pattern that more "concrete items" (clearly defined and less abstract) such as "Cooperation with co-therapist" (item 9) and "Stop and rewind" (item 18; which also had low occurrence), had high reliabilities despite low frequencies. Item 9 can serve as example of an intervention easy to pinpoint, for example, if the rater notices some open communication between the therapists, this counts as an intervention. Items 14 and 16, unlike items 9 and 18, were more difficult to evaluate for quality as well as adherence.

Other items that were difficult to rate for both adherence and quality, and thus deserve careful attention, were "Stimulating

discussions on group norms" (item 8) and "Exploration, curiosity and not-knowing stance" (item 11). Item 11 was used frequently, but it covers a wide range of interventions. The most central aspect of item 11 is to determine whether an open and curious question addresses mental states or not. For example, the intervention "When did he tell you that?" is not aimed at a mental state per se, but depending on the context, some raters may decide to count this as adherence to item 11 – for example if the question makes the patient rethink what really happened, and whether s/he wrongly perceived another person's mental state due to the timing of an utterance. It is difficult to define a clear cut-off without losing some of the flexibility crucial for attuned responsiveness.

We know from previous ratings of non-MBT psychotherapy sessions that non-MBT therapists might display high adherence on items such as "Exploration, curiosity and not-knowing stance" (item 11) and "Affect focus" (item 17). However, the way these therapists intervene is most often different from an MBT approach. They often receive a low quality rating, and raters might be bewildered by boundary occurrences (interventions that border on what might be called MBT). For item 11, the eight PDG sessions had a mean adherence rating on 14 (number of observed interventions), and three for quality. The eight MBT-G sessions had a mean adherence rating of 5, and a mean quality score of 5. The high frequency of low quality item 11-interventions in the rated PDG sessions may account for some of the observed difficulty in rating this item. Still, the manual should be more specific with respect to what counts as adherence and high versus low quality for this item.

From a psychometric perspective, items with low occurrence (e.g., items 9, 14, 16 and 18) may be seen as redundant. However, as underlined in the manual, these items are essential ingredients in a larger treatment "potion:" "The unique aspect of MBT lies less in each individual item per se, than in the overall 'package' of item design and context" (Karterud & Bateman, 2010, p. 26). The robust reliability of the overall ratings indicates that raters manage to capture (agree on) the overall flavor of MBT, even if they disagree on certain items.

Two items that proved difficult to rate were adherence for "Handling psychic equivalence" (item 16), and quality for "Handling pretend mode" (item 15). These two items are both central to the overall theory of mentalization and MBT. For item 16, the 8 PDG sessions had a mean adherence rating of 0, and 2 for quality. The 8 MBT-G sessions had a mean adherence rating of 1, and a mean quality score of 3. Item 15 is not rated for adherence, but both the PDG and MBT-G sessions had a mean quality rating of 2. Both items displayed low variance, high residual variance, and low reliability. In this case, it is unclear whether the group therapists delivered interventions for item 15 and 16 which were poor and/or unclear, or if the concepts of pretend mode and psychic equivalence were somewhat unclear for both therapists and raters. However, taking the small variance into account, the reliability is rather good for these items. Items 15 and 16 should be object for more research, and the manual made more "concrete" for both items.

Limitations

The generalizability of our findings is restricted by several limitations. Firstly, as mentioned above, the raters were not blind

to treatment modality (PDG or MBT-G), and this could have influenced the reliability. However, there were only minor differences between the two modalities and the combined reliability. In the current study, two therapists were rated four times (both in PDG and MBT-G). We cannot exclude the possibility that repeated ratings of the same therapists may have artificially increased inter-rater conformity. Thus, future studies should apply these scales to larger samples of both patients and therapists.

Utility

The MBT-G-AQS may contribute to future psychotherapy research by assuring internal validity and contribute to research on adherence and quality as possible moderators and mediators of treatment outcome. The scale can additionally be used for training and clinical purposes: assessing and providing feedback about therapeutic quality and adherence enables therapists and supervisors to stay on course.

CONCLUSION

The current results demonstrate that the MBT-G adherence and quality scale is a reliable instrument for rating adherence to and quality of mentalization-based group therapy with as few as two raters for the entire scale, and with one rater for overall/global assessment of MBT-G. Some items, especially "Handling pretend mode" and "Handling psychic equivalence" need more empirical attention, as our results indicate these items to be inadequately defined and understood. The scale can be applied for quality assurance, training, and supervision.

REFERENCES

Arnevik, E., Wilberg, T., Urnes, O., Johansen, M., Monsen, J. & Karterud, S. (2009). Psychotherapy for personality disorders: Short term day hospital psychotherapy versus outpatient individual therapy – A randomized controlled study. *European Psychiatry*, 24, 71–78.

Bales, D. L., Timman, R., Andrea, H., Busschbach, J. J., Verheul, R. & Kamphuis, J. H. (2015). Effectiveness of day hospital mentalization-based Ttreatment for patients with severe borderline personality disorder: A matched control study. *Clinical Psychology & Psychotherapy*, 22, 409–417.

Barber, J. P., Liese, B. S. & Abrams, M. J. (2003). Development of the Cognitive Therapy Adherence and Competence Scale. *Psychotherapy Research*, 13, 205–221.

Bateman, A. & Fonagy, P. (2001). Treatment of borderline personality disorder with psychoanalytically oriented partial hospitalization: An 18-month follow-up. *American Journal of Psychiatry*, 158, 36–42.

Bateman, A. & Fonagy, P. (2009). Randomized controlled trial of outpatient mentalization-based treatment versus structured clinical management for borderline personality disorder. *American Journal of Psychiatry*, 166, 1355–1364.

Bateman, A. & Fonagy, P. (2016). *Mentalization based treatment for personality disorders: A practical guide*. Oxford: Oxford University Press.

Burlingame, G. M., MacKenzie, K. R. & Strauss, B. (2004). Small group treatment: Evidence for effectiveness and mechanisms of change. *Handbook of psychotherapy and behavior change*, 5, 647–696.

Cardinet, J., Johnson, S. & Pini, G. (2010). *Applying generalizability theory using EduG*. New York: Routledge.

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

Gutermann, J., Schreiber, F., Matulis, S., Stangier, U., Rosner, R. & Steil, R. (2015). Therapeutic adherence and quality scales for Developmentally Adapted Cognitive Processing Therapy for adolescents with PTSD. *European Journal of Psychotraumatology*, 6, 26632.

Hagtvet, K. A. (1997). The function of indicators and errors in construct measures: An application of generalizability theory. *Journal of Vocational Education Research*, 22, 247–266.

Karterud, S. (2012). *Manual for mentaliseringsbasert gruppeterapi (MBT-G)*. Oslo: Gyldendal.

Karterud, S. (2015). *Mentalization-based group therapy (MBT-G): A theoretical, clinical and research manual*. London: Oxford University Press.

Karterud, S. & Bateman, A. (2010). *Mentaliseringsbaseret terapi. Manual og vurderingsskala*. Oslo: Gyldendal.

Karterud, S., Pedersen, G., Engen, M., Johansen, M. S., Johansson, P. N., Schlüter, C. *Et al.* (2013). The MBT adherence and quality scale (MBT-ACS): Development, structure and reliability. *Psychotherapy Research*, 23, 705–717.

Kvarstein, E. H., Pedersen, G., Urnes, Ø., Hummelen, B., Wilberg, T. & Karterud, S. (2015). Changing from a traditional psychodynamic treatment programme to mentalization-based treatment for patients with borderline personality disorder – does it make a difference? *Psychology and Psychotherapy: Theory, Research and Practice*, 88, 71–86.

McGlinchey, J. & Dobson, K. S. (2003). Treatment fidelity assessment in cognitive behavioral therapy. *Journal of Cognitive Psychotherapy: An International Quarterly*, 17, 299–318.

Möller, C., Karlgren, L., Sandell, A., Falkenström, F. & Philips, B. (2016). Mentalization-based therapy adherence and competence stimulates in-session mentalization in psychotherapy for borderline personality disorder with co-morbid substance dependence. *Psychotherapy Research*, https://doi.org/10.1080/10503307.2016.1158433.

Perepletchikova, F., Hilt, L. M., Chereji, E. & Kazdin, A. E. (2009). Barriers to implementing treatment integrity procedures: Survey of treatment outcome researchers. *Journal of Consulting and Clinical Psychology*, 77, 212–218.

Perepletchikova, F., Treat, T. A. & Kazdin, A. E. (2007). Treatment integrity in psychotherapy research: Analysis of the studies and examination of the associated factors. *Journal of Consulting and Clinical Psychology*, 75, 829–841.

Rossouw, T. I. & Fonagy, P. (2012). Mentalization-based treatment for self-harm in adolescents: Arandomized controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, 51, 1304–1313.

Shavelson, R. J. & Webb, N. M. (1991). *Generalizability theory. A primer*. Newbury Park, CA: Sage.

Shavelson, R. J., Webb, N. M. & Rowley, G. L. (1989). Generalizability theory. *American Psychologist*, 44, 922–932.

Stoffers, J. M., Völlm, B. A., Rücker, G., Timmer, A., Huband, N. & Lieb, K. (2012). Psychological therapies for people with borderline personality disorder. *Cochrane Database of Systematic Reviews*, 8, https://doi.org/10.1002/14651858.CD005652.pub2.

Swiss Society for Research in Education Working Group (2010). *EDUG user guide*. Neuchatel: IRDP.

von Consbruch, K., Clark, D. M. & Stangier, U. (2012). Assessing therapeutic competence in cognitive therapy for social phobia: Psychometric properties of the Cognitive Therapy Competence Scale

for Social Phobia (CTCS-SP). *Behavioral and cognitive psychotherapy*, 40, 149–161.

Wampold, B. E. & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work*. New York, NY: Routledge.

Waltz, J., Addis, M. E., Koerner, K. & Jacobson, N. S. (1993). Testing the integrity of a psychotherapy protocol: assessment of adherence and competence. *Journal of Consulting and Clinical Psychology*, 61, 620–630.

Weck, F., Weigel, M., Richtberg, S. & Stangier, U. (2011). Reliability of adherence and competence assessment in psychoeducational treatment: Influence of clinical experience. *The Journal of Nervous and Mental Disease*, 199, 983–986.

Yalom, I. D. & Leszcz, M. (2005). *Theory and practice of group psychotherapy*. New York: Basic books.

APPENDIX

Rating scale for Mentalization-based Group Therapy

Rater ___ Rating date _____ Therapists _____ Group _____ Session date _____
 Overall rating of MBT adherence _____ MBT quality _____
 Running notes:

Item name	Adherence	Quality
1. Managing group boundaries		
2. Regulating group phases		
3. Initiating and fulfilling turntaking		
4. Engaging group members in mentalizing external events		
5. Identifying and mentalizing events in the group		
6. Care for the group and its members	No rating	
7. Managing authority	No rating	
8. Stimulating discussions on group norms		
9. Cooperation with co-therapist		
10. Engagement, interest and warmth	No rating	
11. Exploration, curiosity and not-knowing stance		
12. Challenging unwarranted beliefs		
13. Regulating emotional arousal	No rating	
14. Acknowleding good mentalizing		
15. Handling pretend mode	No rating	
16. Handling psychic equivalence		
17. Affect focus		
18. Stop and rewind		
19. Focus on the therapist – patient relationship		

Rating scale for Mentalization-Based Group Therapy quality

This is a table used for rating therapist's interventions during group therapy. The table describes the quality level 4 ("good enough"). For more detailed descriptions we refer to the manual.

Item name	Quality level 4 («good enough»)
1. Managing group boundaries	The group is functioning smoothly with respect to boundary issues. The therapists identify boundary relevant events and comment and deal with them in ways which seem appropriate and clarifying for the group as a whole.
2. Regulating group phases	At least two phases are addressed in a way that engages members to reflect upon the possibilities and choices they have.
3. Initiating and fulfilling turntaking	The therapists themselves take initiative and they also follow up patients' initiatives for turntaking. They contribute to the unfolding of the story and identification of relevant scenes, intervene in ways that facilitate a comprehensive narrative and keep a focus on emotions, mental states and interpersonal interactions.
4. Engaging group members in mentalizing external events	The therapists invite the other group members, implicitly or explicitly to clarify relevant events and engage members to participate in a collective exploration of the mental states involved therein.
5. Identifying and mentalizing events in the group	The therapists identify some important events in the group and engage group members in a collective exploration which seems meaningful and clarifying.

(continued)

Table (continued)

Item name	Quality level 4 («good enough»)
6. Care for the group and its members	At this level, the group process is on the even when it comes to care. The therapists seem to have an awareness regarding negative comments between group members and are quick to intervene in such situations.
7. Managing authority	The therapists seem calm and confident as MBT-G therapists. In theory and practise they stand up for the group's basic values.
8. Stimulating discussions on group norms	The therapists take initiative to norm discussions, engage in an interested way in spontaneous discussions and try to modify restrictive group solutions which are being made, if these are not challenged by other group members.
9. Cooperation with cotherapist	There seems to be a confident relationship between the therapists, their interventions are complimentary, and they communicate with each other with open, reflective comments.
10. Engagement, interest and warmth	The therapists appear genuinely warm and interested in each member and the group as a whole. The rater gets the impression that the therapists care in a positive way. Several interventions and their stance indicate this.
11. Exploration, curiosity and not-knowing stance	The therapists pose appropriate questions designed to promote exploration of the patients' and other's mental states, motives and emotions and communicate a genuine interest in finding out more about them.
12. Challenging unwarranted beliefs	The therapists confront and challenge unwarranted opinions about oneself or others in an appropriate manner.
13. Regulating emotional arousal	The therapists play an active role in terms of maintaining emotional arousal at an optimal level (not too high so that patients lose their ability to mentalize and not too low so that the session becomes meaningless emotionally).
14. Acknowleding good mentalizing	The therapists identify and explore good mentalizing and this is accompanied by approving words or judicious praise.
15. Handling pretend mode	The therapists identify pretend mode sequences and intervene to improve mentalizing capacity.
16. Handling psychic equivalence	The therapists identify psychic equivalence functioning and intervenes to improve mentalizing capacity.
17. Affect focus	The interventions focus primarily on emotions – more than on behavior. The attention is particularly directed at emotions as they are expressed in the here and now in the group, and particularly in terms of the relationship between patients and between patients and therapists.
18. Stop and rewind	The therapists identify at least one incident in which patients describe interpersonal events in a non-coherent and affected way, tries to slow down the pace and find out about the event step-by-step. In a similar way, the therapists halt events in the group that tend to be destructive and take initiative to explore the sequence together with the patients.
19. Focus on the therapist – patient relationship	The therapists comment on and attempt to explore, together with the patients, how the patients relate to the therapist during the session and stimulate reflections on alternative perspectives whenever appropriate. The therapists speak about their own feelings and thoughts, related to the patients, and by this they try to engage all parties in mutual exploration.

II

ORIGINAL PAPER

Battles of the Comfort Zone: Modelling Therapeutic Strategy, Alliance, and Epistemic Trust—A Qualitative Study of Mentalization-Based Therapy for Borderline Personality Disorder

E. J. Folmo[1] · S. W. Karterud[2] · M. T. Kongerslev[3,4] · E. H. Kvarstein[5,6] · E. Stänicke[7]

Abstract

We propose a model for how therapeutic strategy, alliance, and epistemic trust interact to foster or hinder therapeutic processes. Four individual mentalization-based treatment (MBT) sessions were subjected to an in-depth qualitative comparison and interpretative phenomenological analysis. Two sessions had high adherence and quality ratings, and two exemplified low evaluations. The sessions were from an MBT program for patients with borderline personality disorder. The high-rated therapists were more prone to strategically identify and investigate maladaptive patterns, were more challenging, and brought the patients out of their comfort zone. This therapeutic endeavour seemed to facilitate therapeutic alliance and a productive therapeutic process. Low-rated therapists seemed to be brought out of their own comfort zone (e.g. transferences/counter-transferences), and attempted to amend the relational atmosphere by being supportive. In these sessions, the therapeutic alliance seemed weak, and therapeutic progress was not observed. When therapists strategically and competently challenged problematic patterns, despite disclosing discomfort, alliance was strengthened. It seemed that a clear therapeutic strategy, and skilfull battling of the patients' comfort zone, fostered the therapeutic process. We hypothesize that epistemic trust may develop as a product of a fruitful and persistent focus on tasks and goals in therapy.

Keywords Mentalization-based treatment (MBT) · Interpretative phenomenological analysis (IPA) · Strategic competence · Therapeutic alliance · Process research

Introduction

Mentalization refers to the ability to understand and interpret behaviours of self and others as expressions of intentional mental states such as feelings, wishes, goals, desires or needs (Fonagy et al. 2002). It develops from early infancy, through attachment relationships and care. The attachment figure is a source for physical security, emotional support, mental attention, knowledge, and culture. Recently, the concept of *epistemic trust* (Fonagy et al. 2018) was introduced to explain the relation between attachment and mentalizing. An attitude of epistemic trust, in contrast to *epistemic freezing*, implies that the listener is ready to take in personally relevant knowledge about the social world. The concepts of mentalization and more recently, epistemic trust, have particularly been advocated in treatment of borderline personality disorder (BPD). The field of psychotherapy research lacks narratives of the phenomenology of different core components and how they may work together. In the present qualitative study of BPD therapy sessions displaying very

✉ E. J. Folmo
  espfol@ous-hf.no

1   Norwegian National Advisory Unit on Personality
    Psychiatry, Section for Personality Psychiatry, Oslo
    University Hospital, Ullevaal, Nydalen, PO Box 4956,
    0424 Oslo, Norway

2   The Norwegian Institute for Mentalizing, Oslo, Norway

3   Psychiatric Clinic Roskilde, Region Zealand Psychiatry,
    Roskilde, Denmark

4   Department of Psychology, University of Southern Denmark,
    Odense, Denmark

5   Section for Personality Psychiatry, Oslo University Hospital,
    Oslo, Norway

6   Institute of Clinical Medicine, University of Oslo, Oslo,
    Norway

7   Department of Psychology, University of Oslo, Oslo, Norway

high and low ratings of adherence and competence, we aim to elaborate on aspects of therapist strategy, alliance, and epistemic trust.

Borderline Personality Disorder and Specifically Tailored Psychotherapy

Patients with borderline personality disorder (BPD) are characterized by insecurity in close attachment relationships, problems of emotional regulation, and a reduced ability to mentalize (Bo et al. 2017). Currently, there are eight specific, evidence-based treatments for BPD (Stoffers et al. 2012). These treatments are all extensive, highly structured, and target core aspects of BPD. One of these is mentalization-based treatment (MBT). Its efficiency for BPD is established in several studies, of which three are randomised controlled trials from UK (Bateman and Fonagy 2001, 2009; Rossouw and Fonagy 2012), and two are naturalistic comparisons replicating positive results in settings outside UK (Bales et al. 2015; Kvarstein et al. 2015). Treatment manuals specifying the style of intervention and reliable integrity measures for therapist interventions exist for both the individual (Karterud et al. 2013) and group components (Folmo et al. 2017).

The Impact of Therapeutic Alliance Across Specific Approaches

Research focusing on mechanisms of change in psychotherapy, has emphasized qualities of the therapist-patient dyad. A therapist's ability to form and maintain a therapeutic alliance (goals, tasks, and personal bond; Bordin 1979) is reckoned as a robust predictor of outcome in psychotherapy. It is known to predict more variance in outcome than the application of a technique, strategy or (*bona fide*) treatment approach alone (Wampold and Imel 2015). However, the process and outcomes of therapy are a result of a complex interplay between therapeutic factors, and specific types of therapy may differ in their involvement and dependence of aspects of alliance (Nissen-Lie et al. 2015). The therapeutic dyad clearly also depends on the patient's ability to form a personal bond to the therapist, create goals and understand the mutual tasks of therapy. Typical aspects of the relational problems in BPD are hostility, insecure attachment, and disturbed epistemic trust (Bo et al. 2017). These are factors which may severly challenge the therapeutic alliance. It is of interest to understand how a therapeutic alliance can be formed and fostered in such circumstances.

Therapeutic Alliance and Clinical Expertise

The mere "relationship" with a therapist is, in itself, insufficient (Laska et al. 2014) for positive outcome, and

therapeutic competence has considerable relevance. For unknown reasons, some therapists seem able to nurture and negotiate therapeutic alliances significantly better than others (Lemma et al. 2011). Across therapy approaches, therapists will apply "strategic competence" (Killingmo et al. 2014) to navigate and structure sessions. We understand strategic competence as the totality of the therapist's understanding of psychotherapy, knowledge of the diagnosis and the patient, and the specific relation. Rønnestad (2016) identifies a combination of a deep engagement in the client's welfare, together with a willingness and capacity to confront the client's dysfunctional behaviour as one of six important characteristics of clinical expertise. In treatment of poorly functioning patients with BPD a willingness to confront maladaptive patterns, may be crucial. However, such confrontation is challenging for both therapist and patient, may represent an interpersonal or emotional "battle of the comfort zone", and needs to be managed with care.

Therapeutic Alliance Challenged by Countertranference

Countertransference reactions may be of particular importance in psychotherapy for BPD (Betan et al. 2005), and are also relevant in structured therapies, such as MBT (Morken et al. 2014). Negative countertransferences in therapists can include feeling helpless, overwhelmed or overinvolved (Colli et al. 2014). Rønnestad (2016) has indeed called for more in-depth investigations of treatments with "difficult to treat clients". Specifically structured treatments aim to represent helpfull strategies in the management of poorly functioning patients. The specified model may then serve as a potential vehicle for the therapeutic alliance.

Therapeutic Alliance and Therapist Model Fidelity

In an MBT study of BPD patients with substance abuse, Möller et al. (2017) reported that high therapist fidelity was associated with an increase in the patients' reflective functioning (operationalization of mentalization; Fonagy et al. 2002) during therapy sessions. In this case, high competence in MBT was seen to induce a productive process of change in core pathology. Nevertheless, little research has focused on how the therapists in evidence-based treatments tailor the specific technique to the patient; how therapists using a certain method, may facilitate alliance and epistemic trust. Hence, there is a pressing call to investigate how (skilled) therapists adapt their specific therapeutic method to the individual patient and thus, integrate the potentially conflicting perspectives—specific treatments and common factors approaches (Laska et al. 2014).

The Present Study

The present study is a qualitative analysis aiming to explore therapeutic dialogues in therapy sessions in light of therapists' strategic competence, patients' indication of epistemic trust and the collaborative therapeutic alliance. For this purpose, we investigated the specific approach, MBT, as a specific treatment for poorly functioning patients with BPD. We selected therapy sessions with high and low ratings of MBT treatment fidelity (Karterud et al. 2013). In studying the transcripts, we sought to understand what influenced the therapists in the sessions, how they maneuvered the topics, how they handled difficult emotions, possible transferences and countertransferences, and the strength of the therapeutic alliance. The results of the qualitative analysis led us to suggest a model of the interaction between these different aspects—alliance, strategy, and epistemic trust.

Materials and Methods

Sessions were selected by purposeful sampling (Patton 1990). The four most deviant (extreme) sessions were sampled from a total of 108 individual MBT sessions assessed with the fidelity scale for MBT-I (Karterud et al. 2013). Ratings were done as a regular, quality ensurance service procedure provided by the Quality Lab for Psychotherapy at Oslo University Hospital, Norway (http://www.mbt-lab.no). The authors reached consensus after independent ratings of the sessions. Rater reliability (estimated on the basis of 30 fidelity ratings) was high (mean value, absolute G coefficients, adherence: 0.95, quality: 0.90). Two authors in this paper (EF and SK) were raters.

The fidelity ratings include MBT adherence and quality. Adherence ratings count the interventions compliant with the 17 items of the fidelity measure. Quality is assessed for each identified item on a 1–7 Likert scale. In addition, global adherence and quality scores are decided for the session as a whole (overall clinical judgement). The cut-off for acceptable MBT-fidelity is four or above. MBT interventions are predominantly characterized by a clear focus on exploration of mental states.

The investigated sessions were all part of MBT programs. Two sessions with high MBT ratings (Adherence: 7; Quality: 7), and two with low ratings (2/2) were selected from Norwegian, Danish and Swedish MBT teams. At the time of video-recordings, treatments had lasted various lengths of time (range 6–24 months). The four therapists were affiliated within MBT teams, were experienced psychotherapists, had advanced MBT training, and received regular MBT supervision. Therapist age-range: 37–65 years. Standard MBT includes patients with personality disorders and core BPD pathology and combine individual and group therapy, emphasize treatment formulations and initial psychoeducation (Karterud 2012; Karterud and Bateman 2010).

For the qualitative process studies, video recordings of the selected four sessions were transcribed, and personal data anonymized. Patients and therapists gave their written, informed consent to participate in the project. The study was approved by the Privacy Ombudsman at Oslo University Hospital.

Qualitative Data Analysis

Our intention was to investigate the phenomena beyond concepts that are defined and operationalized in existing literature. We chose interpretative phenomenological analysis (IPA; Smith et al. 2009) as it allows a fundamental investigation of phenomena like alliance and strategic competence, and has been employed in a number of papers in clinical and counselling psychology (e.g., Østlie et al. 2016; Smith 2011). The transcripts were analysed according to the IPA framework (Smith et al. 2009) in five steps:

(1) The four sessions were transcribed and studied in detail, and discussed in depth, in order to include as many viewpoints as possible (therapist, patient, overarching, synthesis). During this process the first author was in contact with all other authors, discussing transcripts in-depth with the second (SK) and fourth author (EK).

(2) The first, fourth, and last author (EF, EK and ES) sought to phenomenologically investigate the therapeutic alliance (goals, tasks, and personal bond). Agreement on goals could be identified by indications of a mutual idea of achieving improvement. Agreement on tasks was interpreted from the patient's willingness to engage in therapy, participate in a mentalizing discourse or identify, accept and process problematic themes and behaviour patterns. The personal bond could be deduced by patient expressions indicating confidence in the therapist being able to help (aspect of epistemic trust) and a degree of genuine relating, e.g., the patients' trust that the therapist really cared and understood.

(3) Emergent themes identified by (EF) were frequently discussed with the second (SK), fourth (EK) and last author (ES). We looked for possible sequential patterns, how interventions were timed, and identifiable strategies.

(4) The first (EF), fourth (EK), and last author (ES) employed different theories and concepts (e.g., alliance, common factors, strategic competence, MBT, psychoanalytic theory, attachment theory) to illuminate the perceived patterns.

(5) In a final discussion, on the basis of steps 1–5: The first (EF), fourth (EK), third (MK) and last author (ES)

decided on the major recurrent themes/patterns in the sessions.

Results

In the selected sample, the high-rated sessions were characterized by stable focus on mental states (mentalization). The interventions built logically on each other and seemed guided by an overarching strategy: If one intervention failed, the therapists pursued the same goal by another route. In the low-rated sessions, interventions were more seldom, and often lacked a clearly detectable plan or overarching pattern. The high rated sessions were characterized by the therapists being more mentally involved, more active. They also seemed able to manage their own countertransference, focus on affects, keep a mentalizing focus, and challenge the patient in an emphatic and transparent way. In particular, it seemed that the ability to tolerate negative feelings and bring up difficult themes with the patient distinguished high-rated from low-rated sessions. It seemed that high MBT fidelity implied therapies with more willingness for confrontation, and as such, a willingness from both therapist and patient to move beyond a perceivable "comfort zone". Three major recurrent themes/patterns were thus identified: (1) Alliance; (2) Strategic competence; and (3) Battles of the comfort zone. Therapeutic alliance seemed to be fostered by both strategic competence and battles of the comfort zone.

Theme 1: *Therapeutic alliance*. "Where are we headed? Do we cooperate?" Our first identified theme was well defined by Bordin's therapeutic alliance concept (goals, tasks, and personal bond; 1979). In MBT, the overall aim of therapy is to increase the patient's ability to mentalize. From the therapists point of view, the tasks in a therapy sessions is to maintain a focus on mental states, promote a mentalizing dialogue, and explore mentalizing deficits. From the patients point of view, tasks are to bring in, and be willing to explore, personal issues within a mentalizing framework. A strong alliance indicates that the patient understands that increased mentalizing is the ultimate goal, that s/he agrees to work towards this aim, and believes that the therapist can facilitate this process.

Theme 2: *Strategic competence*. "Given this patient, the goal, situation, and relation, how do we best bring about change?" Strategic competence provides the therapist the broader roadmap of how to navigate, adjust, and tailor the MBT technique to the unique patient, relation, and situation. Strategic competence partially overlaps with the quality score of MBT—it includes the timing, precision and relevance of the interventions. Skillful application of MBT includes an overarching ability to navigate (strategic competence) not defined by the MBT manuals.

Theme 3: *Battles of the comfort zone*. "How do we stay on course? Can we challenge maladaptive patterns?" The application of a specific technique, keeping it tailored to the patient, goal, situation, and relation, was a challenge for all therapists. The theme termed "Battles of the comfort zone" emerged when assessing therapist's effort to sustain strategic competence.

Battles of the comfort zone were twofold. From the *therapist perspective*, the persistence of a mentalizing focus, was in some respects, a struggle against resigning to a perhaps, more "comfortable zone", avoiding confrontation (e.g., merely providing supportive therapy). The strong impact of the patient's current mental states such as anger, pretend mode (losing the emotional grounding), teleology (taking actions as evidence for inner states), psychic equivalence (taking own convictions for reality), and possibly also the therapist's own wish for "good transferences", seemed to undermine the application of a focused technique and overall strategy. Battles of the comfort zone also include a *patient perspective*. In high-rated MBT sessions, patients maladaptive behaviors, ways of thinking or relating could be identified and confronted. Avoidance of such confrontation might be to let the patient reside within a (maladaptive) comfort zone. In low rated MBT sessions, the main therapeutic project (theme 1) was abandoned, and these sessions did not reveal relevant MBT therapeutic work. However, in a successful, and repeated confrontative process, as illustrated in the high-rated sessions, the alliance not only endured the strain, but even seemed strengthened by the mutual effort. Our two first identified themes (alliance and strategic competence) seemed to work together and result in beneficial therapeutic work.

Four Case Examples

Below we present our analysis of the three themes in the sessions.

Diane and Her Therapist: Losing Authority and Losing Battles

Diane was a woman in her late 20 s. Her therapeutic project (in the session) was not clear, and she displayed a wide repertoire of strategies to avoid working on her problems in therapy. By attacking, putting down, refuting, appealing to, rejecting, and directly contradicting her therapist, she focused her narrative on several themes, mostly in a pseudo-mentalizing way. She blamed others and her life-situation for her problems, and wanted the therapist to support this view.

Diane opened the session by inquiring whether the therapist had sent a health statement on her behalf: *"Yes. Did you send the statement?"* Her tone was harsh and judgemental. When the therapist turned defensive and uncertain, she

immediately followed up by saying: *"It should have been sent two weeks ago"*, in a way which indicated frustration with the therapist. Next, Diane confronted the therapist for mislabelling her feeling of anger in the previous session: *"Last session, I got angry with you. You said I was irritated, but I wasn't, I was angry!"* The therapist misunderstood her, laughed, and again underestimated her feelings. Diane moved on to say that her problems stemmed from other people, and not from herself. After a while, the therapist vaguely suggested that the patient's views were not necessarily the only reality. Diane immediately refuted this perspective, stating that she took no responsibility for her problems: *"… you made it only my experience and not an actual reality… then you are kind of placing responsibility on me for a situation that is really not my responsibility."* At the end of the session the therapist offered Diane an extra session. Diane turned down this offer, saying that it would not help.

Alliance The patient exhibited little confidence in the therapist and statements explicitly demonstrated a lack of alliance. The emotional level was high. Diane was not able to understand or consider most of the therapist interventions. Interventions did not address the actual relationship or therapeutic project (alliance level). In this case, the possibility for battles of the comfort zone were lost on the alliance level.

Strategic Competence The therapist's initial attempt to laugh away the theme of the patient being "angry and not irritated with him/her" was out of tune with Diane, and the entire session was coloured by a lack of therapists' direction, authority and clarity. Interventions were vague, often only initiated, but not followed up. Possible therapist strategies were outmaneuvered. The therapist missed several opportunities to explore how Diane's statements made sense, or confront non-mentalizing. The most frequent intervention was *"Ehm"*, suggesting an attempt to be warm and supportive. Increasingly, the therapist seemed to strive for a pleasant climate (which often resulted in an even lower interpersonal temperature). At one crucial moment, Diane displayed personal vulnerability in a relational context, but at that point, the therapist missed the invitation to explore mental content, and instead pursued a concrete detail. Diane: *"Ehm… Because… I really felt that I wasn't… seen, in a way, at all. By her. Ehm… Therapist: When did you…?"* Diane: *"Saturday. Therapist: Saturday, OK. Yes, you said that. Yes"*.

Battles of the Comfort Zone Early in the session, the therapist seemed outplayed by their own countertransference (e.g., feeling overwhelmed, helpless, and fearing Diane's anger) and the therapists mentalizing capacity seemed effected. Less able to guide, challenge or question the patient's mental states, the therapist gradually retreated to a supportive and submissive stance. The therapist attempted to challenge Diane when she talked about the other students at her school being the cause of her problems: *"Mm. You kind of.. yes. Because what I was interested in understanding, was*

something like why it is a bigger problem for you than for others, that is what was maybe… that is what was…"*. However, in response to Dianes confronting style, the therapist gradually started to excuse him/herself for questioning her position: *"Yes. No, I was also thinking… it wasn't right… it was foolish to say that… negative attitude and that, so… but I still think that, OK, maybe other people have different…"* Towards the end of the session, Diane said she really needed to finish a paper over the next few days. The therapist then suggested that they should have kept the content of the session more superficial. Diane strongly rejected this argument, leaving the therapist bewildered, still out of touch. Therapist: *"We could have kept it a bit superficial here, but…"* Diane: *What's the point of that? Therapist: Yes, what's the point of that. Right. More superficial or… More focused on the concrete, or… yes. I think it was very important that we spoke about this.."*. The dialogue in this session, indicates that Diane was winning a battle of the comfort zone without resolving her maladaptive, prementalistic, modes of experiencing (pretend mode, psychic equivalence and teleology). Diane: *"This is not something I can do much about. And… I don't see any point in having a positive attitude to something negative."*

Monica and Her Therapist: Protecting the Patient from Therapy

Monica, a woman in her early 20 s, had suffered a violent sexual assault and subsequently missed several sessions. The session was her first since the incident. She conveyed that she lacked energy and did not sleep well. In the session, she seemed uninterested in resuming psychotherapy. The therapist did not challenge the patient. The therapeutic strategy was resigned early in the session. The session included some enquiry, information and continued with a sequence about Monica's wish to buy a new dress. The patient finally wanted to end the session five minutes early, *"as they had nothing important to talk about"*. The therapist agreed.

Alliance Most interventions aimed for a positive personal bond. The relationship or therapeutic project was not addressed directly. Monica had one utterance addressing alliance to the group *"No, I am actually quite excited about getting back there, because it has been pretty much... a lot happening there."* However, she did not seem enthusiastic about the ongoing individual therapy session and took the opportunity to end the session early. The alliance seemed weak.

Strategic Competence Monica's therapist sought a warm, gentle, considerate atmosphere throughout the session, asked practical questions, validated responses, but largely avoided exploration and refrained from challenging the patient. Brief inquiries included details after the assault (had the rapist been caught: *"You don't know, or do you know that he hasn't*

been caught?"; was support from health care and judicial system sufficient), on post-traumatic symptoms (dreams/ nightmares; fear of walking alone in the dark), and functioning (was coming to two group sessions too much at the moment, was she able to continue at school: *"Have you managed to get back on your feet with regards to … school and… or have you…";* how was her social network, *"Who is close by you now?";* and how were other things in her life, e.g., *"What else is happening to you?").* The therapist provided news from the group, advice on sleep medication, and normalized symptoms in light of the recent incident.

Battles of the Comfort Zone The therapist had a strategy of not confronting the patient too much in the current situation—it is unclear what was the patient's perspective as she had difficulties with elaborating on her own mental state. This is captured by the therapist. Therapist: *"But those thoughts that are coming in lots... those thoughts, what are... I would have liked to hear."* Monica: *"Well, this is what I have been telling you".* Therapist: *"Yes. But are there any more?"* Monica: *"No."* Therapist: *"No.. no...?.. content, no kind of depressive... no kind of wish that you were... no kind of...?"* Monica: *"No. I am more kind of indifferent, really."* Therapist: *"Indifferent."* Monica: *"Yes".* Nevertheless, countertransference appear to be present, effecting the quality of the session. The fact that Monica had not turned up to therapy for a while was brought up. However, the question was framed so it could be precieved rather as difficult for the therapist, who had been worried, than care for the patient. The therapist also brought up missed sessions of group therapy, but abandoned the theme when Monica explained her total lack of energy after the traumatic event. The therapist often seemed to lack curiosity for the answers to own questions and in one example, the therapist gave a conclusion on behalf of the patient. Therapist: *""Who is close by you now?"* Patient: *"Right now it is S and Y, family."* Therapist: *"Yes. But you are a little lonely…."* The struggle of the comfort zone in this case seems to end up with a dialogue devoid of any exploration of mental states, both parts avoiding discomfort, which nevertheless seemed to be present. The therapist becomes increasingly careful, avoidant of emotional themes, oversupportive perhaps, and the patient increasingly unmotivated, but possibly, left in a vulnerable state. Implicitly, the therapist may have conveyed compassion, but coupled with possible unresolved countertranferences of helplessness or resignation.

Elsa and Her Therapist: Leaning on the Alliance in the Battle of the Comfort Zone

Elsa was a woman in her early 50 s. She was also a former heroin addict. Recently, she had felt hurt in a group therapy session, and had avoided coming for 4 weeks. This was the most salient subject in the session. The underlying theme of

returning when someone had hurt her was painful for Elsa. She tried several strategies to avoid talking about the group in the session.

Alliance Elsa made seven statements that directly addressed the alliance in highly positive terms. The second one occurred about 10 min into the session: *"Yes, but. Fucking good. How competent you are. Thank you."* From the context, it suggests a genuine sense of being helped (bond part of alliance) and it may indicate an aspect of epistemic trust. One utterance captured some of her inner representation of the therapist's persistent stance: *"Yes but I see, I see what you're saying, I see what you know you see. YES."* Later in the session, Elsa gave a statement concerning the appreciation of new learning: *"It's good that others see things as well, that I don't see."* By the word "others" it is clear in this context that it was the therapist she denoted, although she chooses a less personal and more general phrasing. Elsa's announcement also expresses gratefulness. She recognized her therapist as competent and appreciated his help. In this session the therapeutic dialogue between patient and therapist indicates that the alliance relates closely to patients confidence (experience of new interpersonal learning about herself stemming from the therapy) and enables the therapist to keep a focused strategy.

Strategic Competence The therapist kept a persisting mentalizing stance insisting to talk about Elsa's attendance to group therapy—a part of the MBT program. The therapist's core strategy was close to the MBT manual, with curiosity about mental states, keeping focus on mental content, and being transparent about their own mind. The therapist often started by exploring and clarifying a topic, summarizing or connecting to a larger framework of understanding, and then employing a more challenging stance. For instance, after Elsa had agreed to return to the group, her therapist concluded the theme by highlighting her own responsibility and agency: *"No, and when I asked you about this, it was not to criticize you, but to emphasize the problem with it. There is something that is making it difficult when we talk about it. But the only one who can persuade you to go to the group is you, yourself."* In this session the focused therapeutic strategy seems to relate closely to the therapists specific MBT competence.

Battles of the Comfort Zone The session revealed Elsa's discomfort and her relational issues. She (quite correctly) expected her therapist to challenge her, and tried to avoid such interventions by laughing, distracting and opposing. Elsa's strong appraisals of her therapist could also be interpreted as a defensive strategy, (implicitly) implying that the therapist should be gentle with her, as she was nice to the therapist. However, Elsa's therapist was not led astray by her avoidance strategies. After several interventions, persistently, negotiating a need for talking about the theme, e.g., *"I think we should talk about it now, and then we can*

return to what we were talking about, all right?", the therapist finally succeeded in this first step. In creating this situation the therapist leaned on the therapeutic bond, which seemed good enough to allow the persistence. S/he was then able to say more about why the group is so important for the patient, and how s/he felt somewhat stupid for "nagging about it for the hundredth time", when the patient did not attend the group even though she promised. The following is an example of alliance and strategy working together. The therapist is open about countertransferences. Therapist: *"At the same time, I think like this: Now that we're talking about it, I try in a way, well..it...it is quite difficult, because I can't hide that I think it's good for you to go there. Just because I happen to think so?! But at the same time, I feel that I nag you about this a lot. And then I think like this: Is it because I keep nagging you, that you say yes, that you want to go there? Because you don't go there. And then I feel...well, what am I doing..... and I feel disappointed in a way. We talk about it and you say you will go there and then you don't...."* Elsa and her therapist seemingly agreed on the goals and tasks in the therapy, even though the patient resisted them. In this session, in contrast to the former examples of Diane and Monica, the personal bond (established trust) enabled an explicit battle of the comfort zone, and Elsa, who accepted the struggle, thus achieved a therapeutic focus on her core relational problems. In treatment of patients with severe relational problems, the concept "battles of the comfort zone", depicts a two-way tension within the therapeutic dyad.

Maria and Her Therapist: Using Empathic Focus to Carefully Battle Affect Avoidance

Maria was a woman in her early 30 s. She harboured strong resistance to the conjoint group therapy. When she eventually turned up in the group, she experienced skepticism. This urged her to leave the group. The therapist asked if some of her thoughts and feelings about this could be shared with the group. Maria responded that strangers should have no access to her inner life. This reactivity echoed other relations in her life, and she had lately become rather isolated. The therapist explored various barriers Maria raised in relation to the group in an empathic and steadfast way, which finally allowed Maria's underlying sadness to emerge.

Alliance Maria provided 20 statements concerning alliance. Six of these were connected to a plan of education. If it proved impossible to combine with treatment, she stated that she would choose treatment: *"Yes. Yes, yes, and I am also prepared that, if it should be, that I cannot, so if it should be, that, that my teacher does not want to give me dispensation, then I am fully aware that I will have to drop the education."* We interpreted this statement as reflective of Maris's commitment to the treatment she was receiving. Maria felt diagnostic assessments had been helpful: *"Ehm... but I have only just become really aware of my feelings and my...everything after I got my diagnosis."* Maria indicated that she was not used to be challenged: *"Ehm...so I haven't...I haven't necessarily had to face a lot of...anything in reality."* Inferably, Maria nevertheless, here can be seen to accept this aspect of therapy. An important contributing factor may be that the therapist seemed highly emotionally attuned. Throughout the session she was able to accurately identify the patient's feelings. Consequently, it is likely that the patient felt held and understood in a contingent and congruent way. The treatment was in its beginning, but the alliance already appeared strong.

Strategic Competence The therapist was highly adherent to the manual, had an impressive range of MBT interventions, and awareness of the conjoint therapy aspect. The therapist validated, encouraged, and kept a steadfast focus on mental states throughout the session. This process seemed to stimulate the patient's ability to mentalize others, and facilitated Maria in exploring the experience of the other group members: *"Mm. Do you think that the others notice the feeling you have, that it doesn't concern them?"* and *"What do you think made her say something like that?"* In a playful and gentle way, she further encouraged Maria to mentalize her emotional reactions to the others in the group: *"Did you get a little irritated by her not trying to see it from your perspective... viewpoint? Maybe? The fact that it also could be difficult for them? Do you think that is what made you most irritated?"* The therapist balanced being challenging and supportive, and explored the patient's resistance to the group therapy in great detail, while she semeed to validate Maria's different difficulties in a transparent and clear way: *"Because it, I think, it could also be really difficult to be the new one and kind of have to get in to a group, that already is going, and... try to find one's feet there, and find a place in the group, and I suppose, that too can be really difficult"*. She also normalized Maria's trouble in choosing themes for the group, and actively encouraged her to talk about this in her individual therapy: *"If there is any situation... well how, you could bring something into the group. So we could try to look at that... what could be relevant for you. There are a lot of people who feel like that, that... what exactly do I bring up... what kind of event one should talk about,... that is when you can use our sessions to look at, whether there could be some relevant situations..."*

Battles of the Comfort Zone The main part of the session was spent exploring and gradually challenging Maria's concerns about the group, and reasons for not finding it fruitful. The therapist was steadfast in her focus on mental states and mentalizing of Maria's attachment to and beliefs about the group. This increasingly activated the patient, and resulted in her being "irritated" at the therapist for being "poked". As Maria was brought out of her emotional comfort zone the therapist asked: *"But I'm wondering, what can you notice*

right now, when you are sitting here telling me these things? What are you in contact with now?" Maria said she felt "irritated". The therapist investigated this further by saying: *"So me asking about things, and trying to understand some things, and examining some things together with you, can actually be experienced as irritating?"* Maria confirmed that being "poked" like this by the therapist annoyed her, and then admitted that it was *"not too comf... fantastic"* to say this aloud to the therapist—but she said it with a big smile. It was a relief for Maria to have ventilated her feelings towards the therapist. It seemed to strengthen the bond. Her experience of being different and lonesome filled the last part of the session, now with tears and sadness. She seemingly felt seen, met and held by her therapist and her narrative became more open, personal and in contact with emotions.

Discussion

This qualitative analysis of therapy sessions with high and low-rated MBT fidelity including poorly functioning patients with BPD, highlights interactions between therapeutic alliance and therapists' strategy. We suggest a model where alliance and strategic competence work together, and enable focused, but challenging work with highly sensitive patients and their psychopathology. Further, we postulate that such a process may have the potential of increasing the patient's epistemic trust. A central theme was depicted in the concept "battles of the comfort zone".

Battles of the Comfort Zone: Expanding the Front Line of the Therapeutic Relationship

The low rated MBT sessions highlighted how counter-transferences of being useless, judged/criticized, not knowing enough (incompetent), not being liked, or strong feelings of sympathy, may result in a therapeutic style with too little confrontation. In the low rated sessions, therapists seemed to be avoiding difficult contents or trying to accommodate or please the patient. Therapist interventions included concrete/practical advice or offering extra sessions. The low rated MBT therapists seemed for various reasons to be brought out of *their* comfort zone and their competence was outplayed. These sessions displayed a lack of mentalizing on behalf of the therapist in terms of few MBT interventions and abandonment of the overall therapeutic strategy.

The high rated MBT therapists seemed to have kept their ability for mentalizing during the session, and were able to focus more explicitly on the alliance, and explore possible transference reactions in a transparent manner with the patient. The therapists remained steadfast and committed to the overall goals of trying to increase the patients' mentalizing abilities and seemed to tolerate the patient's anger,

depreciation, abstruseness, or stubbornness as well as the more austere atmosphere that arose when they pursued the patient's problems.

Our analysis suggests that the high rated clinicians were willing to challenge the patients, even though it would temporarily disharmonize the therapeutic relation. High rated therapists identified, investigated, and confronted the patients' problems in a clarifying process, which in turn, further promoted therapeutic alliance. In the low rated sessions, the therapeutic alliance was interpreted as weak, and no positive progress was observed. Low rated therapists were brought out of their own comfort zone (e.g., by transferences and/or counter-transferences), and attempted to amend the atmosphere by being overly agreeable and accommodating.

It seemed that a positive alliance and clear strategic competence were two necessary, coacting components allowing for what we conceptualized as "battles of the comfort zone". The therapist needs a willingness and capacity to confront the client's dysfunctional behaviour (Rønnestad 2016), and a willingness to tolerate the discomfort (e.g., transferences and counter-transferences) this may cause in the session. We propose that, when administered with skill, such "battles of the comfort zone" may evoke an even stronger alliance.

In our sessions, the more there was a sense of genuine warmth (personal bond) in the relation, despite struggles, the more it seemed possible for the therapist to challenge the patient even further. This general sense of a "warm climate", similar to what Sandler (1960) termed background of safety, in the high-rated sessions seemed to enable work on sensitive, but core relational or personal issues. In our analysis, a crucial part of this warmth or background of safety is most accurately seen as trust: It is reasonable that such trust is an accumulated asset built from assimilated experiences of the therapist being able to help.

In the two high-rated sessions, trust evolved through repeated experiences of the therapist being able to guide, reflect, explore, understand, challenge, and/or interpret (help the patient connect specific situations to a larger dysfunctional behavioural pattern) the mental content. It is conceivable that improvement in epistemic trust could evolve from the therapists' willingness to address and confront maladaptive patterns according to an overarching strategy. We postulate that such a process may have the potential of increasing the patient's epistemic trust, which is crucial because therapy then works through three levels. First, the patient's trust in the therapist allows her to learn new content about mental states of self and others. Secondly, the therapy foster mentalization through a process of reflecting mental states. Thirdly, the new content and reflection relaxes a hypervigilance in social situations, which in turn opens for new social learning (Fonagy et al. 2018).

A different conception could be that such battling of the comfort zones induces what Davanloo (1990) refers to as an

"unconscious alliance". This means that the patients' unconscious trust (alliance) is built by the therapists' willingness to directly confront the patients' defences (battle the comfort zones) in order to be helpful. McCullough (1991) found that patients seemed more able to digest the painful information contained in a therapist's confrontation or interpretation when it was paired with a statement that reflected consideration or care—it was detected that confrontations made along with a supportive or empathic statement by the therapist resulted in a greater probability of affective activation.

As we assume that epistemic trust can be gained or regained, the alliance need not be high in all sessions. A treatment may be efficient as a whole, despite some low rated sessions. Consequently, it is more important to negotiate the alliance than to have a positive personal bond at all times (Safran and Muran 2000; Zilcha-Mano et al. 2015).

In the low rated sessions, the patients seemed to command the battles of the comfort zone. In our selection of four sessions, the high rated therapists built on the personal bond and managed to pull the patient towards their common goal. The personal bond appeared as an asset allowing the therapist to challenge the patients' sensitive subjects. The high rated therapists were selective about what s/he wanted to battle (strategic competence). Both Diane and Monica (low rated sessions) displayed low trust in receiving help from their respective therapists. In the session with Monica, the atmosphere was difficult to interpret, her mental state was described as "indifferent", and an increase in mentalizing could not be observed. In the session with Diane, the atmosphere was tense, and the therapist struggled to improve it, but lost focus on the overall therapeutic project in the session. In the high-rated sessions, the general atmosphere was not uncomfortable, but had the distinct quality of the patient both protesting, but gradually working with and accepting challenges. The atmosphere was coloured by the patient's content.

Strengths and Limitations

In line with recommendations for purposeful sampling, we selected the most extreme or deviant sessions in order to illuminate possible themes or patterns (Patton 1990). The logic and power of purposeful sampling lies in selecting information-rich sessions, those from which one can learn a great deal about issues of central importance to the purpose of the research, for in-depth analysis. Hence, our findings depend on the assumption that the four most deviant sessions will inform us about alliance in MBT. One could argue for a larger sample, or for selecting more average sessions.

Smith et al. (2009) underscore that the purpose of IPA is to attempt to gain an insider perspective, while acknowledging that the researcher is the primary analytic instrument. The researcher's beliefs are not seen as biases to be eliminated, but as a necessity for interpretation of the qualitative data. It may thus be regarded a strength that the researchers are experts in the field they investigate (Binder et al. 2012). However, in order to balance possible biases towards MBT, the last author is a psychoanalyst, and had no formal MBT education.

The study focused on aspects of alliance. Alliance may be assessed in a variety of ways, often by quantitative methods such as self-reports, and is shown to predict positive outcome across several measurement methods (Martin et al. 2000). This suggests that trained clinicians should be able to evaluate qualities of therapeutic alliance by observation of in-session processes. Our phenomenological analysis was based on the assumption that alliance could be analysed as the phenomena of the relational process (Henry and Strupp 1994). Built on this fundament, the three aspects of alliance were investigated phenomonologically on the basis of the transcripts. The study is nevertheless limited by a lack of quantitative data which could support our interpretations of alliance.

Conclusion

Based on MBT therapy sessions for poorly functioning patients with BPD, we suggest a model where alliance and strategic competence work together, enabling focused, but challenging work with highly sensitive patients. We postulate that such a process may have the potential of increasing the patient's epistemic trust, which is crucial because therapy then works through the three levels described by Fonagy et al. (2018).

The tension within the therapist-patient dyad was clearly illustrated in all the therapies, challenged therapeutic strategies, and was termed "battles of the comfort zone".

However, within a framework of a trusting alliance, therapists were able to keep a focused strategy and address problems. We suggest that this fruitful interaction, nurtured epistemic trust, and a willingness to manage sensitive topics within the therapeutic dyad. Conversely, poorly demonstrated therapist strategies were coupled with low confidence and lack of alliance in patients, and possibly further enhanced by activation of therapist countertransference. Such interaction implied severely restricted possibility for managing sensitive topics within the therapist-patient dyad. The study raises the question of how not only the bond, but also the task aspect of alliance, may be a crucial factor in treatment of poorly functioning individuals.

Psychiatry, Oslo University Hospital, Norway for providing us with material to this study.

Compliance with Ethical Standards

Conflict of interest None of the authors have any financial disclosure/conflict of interest related to this manuscript.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent: Informed consent was obtained from all individual participants included in the study.

References

Bales, D. L., Timman, R., Andrea, H., Busschbach, J. J., Verheul, R., & Kamphuis, J. H. (2015). Effectiveness of day hospital mentalization-based treatment for patients with severe borderline personality disorder: A matched control study. *Clinical Psychology & Psychotherapy, 22*(5), 409–417.

Bateman, A., & Fonagy, P. (2001). Treatment of borderline personality disorder with psychoanalytically oriented partial hospitalization: An 18-month follow-up. *American Journal of Psychiatry, 158*(1), 36–42.

Bateman, A., & Fonagy, P. (2009). Randomized controlled trial of outpatient mentalization-based treatment versus structured clinical management for borderline personality disorder. *The American Journal of Psychiatry, 166*(12), 1355–1364.

Betan, E., Heim, A. K., Conklin, Z., C., & Westen, D. (2005). Countertransference phenomena and personality pathology in clinical practice: An empirical investigation. *American Journal of Psychiatry, 162*(5), 890–898.

Binder, P. E., Holgersen, H., & Moltu, C. (2012). Staying close and reflexive: An explorative and reflexive approach to qualitative research on psychotherapy. *Nordic Psychology, 64*(2), 103–117.

Bo, S., Sharp, C., Fonagy, P., & Kongerslev, M. (2017). Hypermentalizing, attachment, and epistemic trust in adolescent BPD: Clinical illustrations. *Personality Disorders: Theory, Research, and Treatment, 8*(2), 172–182.

Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research & Practice, 16*(3), 252–260.

Boswell, J. F., Gallagher, M. W., Sauer-Zavala, S. E., Bullis, J., Gorman, J. M., Shear, M. K., … Barlow, D. H. (2013). Patient characteristics and variability in adherence and competence in cognitive-behavioral therapy for panic disorder. *Journal of Consulting and Clinical Psychology, 81*(3), 443–454.

Colli, A., Tanzilli, A., Dimaggio, G., & Lingiardi, V. (2014). Patient personality and therapist response: An empirical investigation. *American Journal of Psychiatry, 171*(1), 102–108.

Davanloo, H. (1990). *Unlocking the unconscious*. Chichester: Wiley.

Folmo, E. J., Karterud, S. W., Bremer, K., Walther, K. L., Kvarstein, E. H., & Pedersen, G. A. (2017). The design of the MBT-G adherence and quality scale. *Scandinavian Journal of Psychology, 58*(4), 341–349.

Fonagy, P., Gergely, G., Jurist, E., & Target, M. (2002). *Affect regulation, mentalization, and the development of the self*. New York: Other Press.

Fonagy, P., Luyten, P., Allison, E., & Campbell, C. (2018). *Reconciling psychoanalytic ideas with attachment theory*. New York: Guilford Press.

Henry, W. P., & Strupp, H. H. (1994). *The therapeutic alliance as interpersonal pro*cess. In A. O. Horvath & L. S. Greenberg (Eds.), *Wiley series on personality processes. The working alliance: Theory, research, and practice* (pp. 51–84). Oxford: John Wiley.

Henry, W. P., Strupp, H. H., Butler, S. F., Schacht, T. E., & Binder, J. L. (1993). Effects of training in time-limited dynamic psychotherapy: Changes in therapist behavior. *Journal of Consulting and Clinical Psychology, 61*(3), 434–440.

Karterud, S. (2012). *Manual for mentaliseringbasert gruppeterapi (MBT-G)*. Oslo: Gyldendal Akademisk.

Karterud, S., & Bateman, A. (2010). *Manual for mentaliseringbasert terapi (MBT) og MBT vurderingsskala. Versjon individualterapi*. Oslo: Gyldendal Akademisk.

Karterud, S., Pedersen, G., Engen, M., Johansen, M. S., Johansson, P. N., Schlüter, C., … & Bateman, A. W. (2013). The MBT Adherence and Competence Scale (MBT-ACS): development, structure and reliability. *Psychotherapy Research, 23*(6), 705–717.

Killingmo, B., Varvin, S., & Strømme, H. (2014). What can we expect from trainee therapists? A study of acquisition of competence in dynamic psychotherapy. *The Scandinavian Psychoanalytic Review, 37*(1), 24–35.

Kvarstein, E. H., Pedersen, G., Urnes, Ø, Hummelen, B., Wilberg, T., & Karterud, S. (2015). Changing from a traditional psychodynamic treatment programme to mentalization-based treatment for patients with borderline personality disorder—Does it make a difference? *Psychology and Psychotherapy: Theory, Research and Practice, 88*(1), 71–86.

Laska, K. M., Gurman, A. S., & Wampold, B. E. (2014). Expanding the lens of evidence-based practice in psychotherapy: A common factors perspective. *Psychotherapy, 51*(4), 467–481.

Lemma, A., Target, M., & Fonagy, P. (2011). *Brief dynamic interpersonal therapy: A clinician's guide*. Oxford: Oxford University Press.

Martin, D. J., Garske, J. P., & Davis, M. K. (2000). Relation of the therapeutic alliance with outcome and other variables: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 68*(3), 438–450.

McCullough, L. (1991). Davanloo's short-term dynamic psychotherapy: A cross-theoretical analysis of change mechanisms. In R. Curtis & G. Stricker (Eds.), *How people change: Inside and outside of psychotherapy* (pp. 59–79). New York: Plenum Press.

Möller, C., Karlgren, L., Sandell, A., Falkenström, F., & Philips, B. (2017). Mentalization-based therapy adherence and competence stimulates in-session mentalization in psychotherapy for borderline personality disorder with co-morbid substance dependence. *Psychotherapy Research, 27*(6), 749–765.

Morken, K., Karterud, S., & Arefjord, N. (2014). Transforming disorganized attachment through mentalization-based treatment. *Journal of Contemporary Psychotherapy, 44*(2), 117–126.

Nissen-Lie, H. A., Havik, O. E., Høglend, P. A., Rønnestad, M. H., & Monsen, J. T. (2015). Patient and therapist perspectives on alliance development: Therapists' practice experiences as predictors. *Clinical Psychology & Psychotherapy, 22*(4), 317–327.

Østlie, K., Stänicke, E., & Haavind, H. (2016). A listening perspective in psychotherapy with suicidal patients: Establishing convergence in therapists and patients private theories on suicidality and cure. *Psychotherapy Research*, 1–14.

Patton, M. Q. (1990). *Qualitative evaluation and research methods*. Thousand Oaks: SAGE.

Rønnestad, M. H. (2016). Is expertise in psychotherapy a useful construct? *Psychotherapy Bulletin, 51*, 11–13.

Rossouw, T. I., & Fonagy, P. (2012). Mentalization-based treatment for self-harm in adolescents: A randomized controlled trial. *Journal of*

the American Academy of Child & Adolescent Psychiatry, 51(12), 1304–1313.

Safran, J. D., & Muran, J. C. (2000). *Negotiating the therapeutic alliance: A relational treatment guide*. New York: Guilford Press.

Sandler, J. (1960). The background of safety. *The International Journal of Psychoanalysis, 41*, 352–356.

Smith, J., Flowers, P., & Larkin, M. (2009). *Interpretative phoneomological analysis: Theory, method and research*. London: SAGE Publications.

Smith, J. A. (2011). Evaluating the contribution of interpretative phenomenological analysis. *Health Psychology Review, 5*(1), 9–27.

Stoffers, J. M., Voellm, B. A., Rücker, G., Timmer, A., Huband, N., & Lieb, K. (2012). Psychological therapies for people with borderline personality disorder. *The Cochrane Library*.

Strupp, H. H. (1998). Negative process: Its impact on research, training, and practice. *Empirical Studies of Psychoanalytic Theories, 8*, 1–26.

Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work*. New York: Routledge.

Zilcha-Mano, S., Solomonov, N., Chui, H., McCarthy, K. S., Barrett, M. S., & Barber, J. P. (2015). Therapist-reported alliance: Is it really a predictor of outcome? *Journal of Counseling Psychology, 62*(4), 568–578.

III

Development of therapeutic alliance in mentalization-based treatment—Goals, Bonds, and Tasks in a specialized treatment for borderline personality disorder

E. J. Folmo , E. Stänicke , M. S. Johansen , G. Pedersen & E. H. Kvarstein

Published online: 11 Oct 2020.

Submit your article to this journal ↗

View related articles ↗

View Crossmark data ↗

Routledge
Taylor & Francis Group

Check for updates

EMPIRICAL PAPER

Development of therapeutic alliance in mentalization-based treatment —Goals, Bonds, and Tasks in a specialized treatment for borderline personality disorder

E. J. FOLMO [1,2], E. STÄNICKE[2], M. S. JOHANSEN [3], G. PEDERSEN[3,4], & E. H. KVARSTEIN[3,5]

[1]*Norwegian National Advisory Unit on Personality Psychiatry, Section for Personality Psychiatry & Specialized Treatments, Division of Mental Health & Addiction, Oslo University Hospital, Oslo, Norway;* [2]*Department of Psychology, University of Oslo, Oslo, Norway;* [3]*Section for Personality Psychiatry & Specialized Treatments, Division of Mental Health & Addiction, Oslo University Hospital, Oslo, Norway;* [4]*NORMENT, KG Jebsen Center for Psychosis Research, Institute of Clinical Medicine, University of Oslo, Oslo, Norway &* [5]*Adult Psychiatry Unit, Institute of Clinical Medicine, University of Oslo, Oslo, Norway*

Abstract
Objective: Mentalization-based treatment (MBT) is an evidence-based long-term treatment for borderline personality disorder (BPD). Alliance is central for effective psychotherapies. Few studies have addressed aspects of working alliance in BPD evidence-based treatments. This study aimed to investigate alliance development in MBT therapies with different clinical outcomes. **Method:** The sample included 155 patients in an MBT programme. Clinical outcomes were based on Global Assessment of Functioning (GAF). The sample was divided in two subgroups according to GAF levels at the end of treatment (cut-off = 60). Working alliance was assessed by patient report (Working Alliance Inventory, subscales, Goals, Bonds and Tasks) and assessed repeatedly over 36 months. The method for statistical analyses was linear mixed models. **Results:** Initial levels of Goals, Bonds, and Tasks did not differ by subgroup, but change over time differed significantly by subgroup. In the good outcome subgroup, ratings of Goals, Bonds, and especially Tasks increased significantly over time. In the poor outcome subgroup, paranoid PD was associated with poorer alliance development over time. **Conclusions:** Good outcome therapies were characterized by a process where the working alliance grew over time. Results encourage an explicit focus on tasks in therapy particularly for patients with high levels of mistrust.

Keywords: mentalization-based treatment (MBT); borderline personality disorder (BPD); working alliance inventory; therapeutic alliance; process research

Clinical or methodological significance of this article: This article points to the clinical importance of maintaining careful alliance work in the treatment of poorly functioning patients with BPD. Such work includes not only a longer-term process of attachment and bonding, but also keeping the goals of therapy understandable, current and updated, and making the therapeutic work, progress, and challenges relevant and explicit. As a specialized treatment for BPD, MBT includes interventions and structure aiming to support therapists and thereby facilitate therapy for poorly functioning patients with considerable emotional and relational problems.

The working alliance predicts approximately 7.5% of the variance in treatment outcomes and is considered a major mechanism of change in psychotherapy (Flückiger et al., 2018; Wampold & Imel, 2015). It has been operationalized in terms of (i) agreement on the *Goals* of treatment, (ii) the formation of personal

Bonds between patient and therapist, and (iii) the therapeutic work process, conceptualized as *Tasks* (Bordin, 1979). Nevertheless, there are still few studies following the development of the working alliance in long-term therapies, especially the subparts of alliance—Goals, Bonds, and Tasks (Stiles & Goldsmith, 2010). The therapeutic alliance may be a crucial factor in treatment for borderline personality disorder (BPD; Dimaggio et al., 2019), but it is poorly investigated in evidence-based treatments for BPD.

The therapeutic alliance is not always easily established in psychotherapy (Colli et al., 2014), not least in the treatment of severe BPD patients. Epistemic trust—that is, the basic ability to trust significant social information from others—is a relevant concept possibly related to the working alliance. It is considered fundamental for the development of interpersonal and relational capacity, and impairment has been linked to BPD (Bo et al., 2017). In psychotherapy, work on difficulties in the alliance may in itself be essential (Safran & Muran, 2000; Wampold & Imel, 2015), and for BPD patients, alliance improvement can even be seen as a treatment outcome (Muran & Barber, 2011). In light of the profound attachment problems apparent among many BPD patients, we see a need for further investigation of alliance processes in structured BPD treatments.

Mentalizing is a core aspect of personality functioning and may be defined as an imaginative mental activity enabling perception and interpretation of mental states (e.g., needs, desires, feelings, beliefs, and goals) (Fonagy et al., 2015). Among patients with BPD, personality problems have been associated with attachment insecurity, tendencies of dysphoria, emotional dysregulation, and social misinterpretations—often in terms of hyper-mentalizing (Sharp, 2014; Vaskinn et al., 2015). Mentalization-based treatment (MBT) is a specialized BPD treatment developed from traditional psychoanalysis and research on attachment and social cognition (Bateman & Fonagy, 2016). Across different treatment theories and techniques, therapists primarily aim to engage the patient in work that feels meaningful, although the emphasis on each alliance component (Goals, Bonds, and Tasks) may be different (Falkenström & Larsson, 2017; Horvath & Greenberg, 1989; Muran & Barber, 2011). In this article we aim to investigate how the different alliance components develop in MBT.

MBT is a long-term, manualized, multicomponent treatment programme (Bateman & Fonagy, 2016; Karterud, 2015). It consists of five structural pillars (Bateman & Fonagy, 2016): (i) specific BPD psychoeducation about core personality problems and the focus of MBT treatment, (ii) an individually adjusted case formulation that is dynamic (changeable during the treatment process), (iii) the combination of individual and group formats of long-term psychotherapy, (iv) a team of therapists working together with regular MBT supervision, and (v) a frame and style of intervention aiming to facilitate a mentalizing process. Together these pillars can be seen as enforcements promoting the development of a therapeutic alliance. The first two explicitly address Goals and Tasks of treatment, the third adds the opportunity for interpersonal bonding (i.e., the bond part of the working alliance), and the fourth and fifth support therapists' interventions, techniques, reflective practice, and handling of countertransference. The latter are considered essentially important as adherent therapist interventions in MBT have been associated with improved reflective functioning (Möller et al., 2017).

Positive effects of MBT are demonstrated in several studies, and outcomes mainly include symptomatic alleviation and the reduction of self-harming or suicidal behaviours and hospital admissions (Volkert et al., 2019). There are yet few investigations of working alliance for patients in MBT. Nonetheless, in a qualitative study of change processes in MBT, Morken et al. (2019) emphasize the importance of repairing alliance ruptures. In other studies, patients' positive experiences include the identification of personality problems, a feeling of symptom improvement, and the content of therapeutic work—learning to regulate oneself, gaining new perspectives, or attending groups (Dyson & Brown, 2016; Johnson et al., 2016; Lonargáin et al., 2017). A recent MBT study pointed to interventions focusing on mentalizing positive affects as possibly beneficial for alliance (Harpøth et al., 2019). Moreover, in a study of MBT group therapy, interpersonal personality features influenced the establishment of a working alliance in the group (Euler et al., 2018), and the authors recommended particular apprehension of BPD patients' relational bias and hyper-mentalization in the early phase of therapy.

Specialized approaches involve specified therapeutic formats and techniques, and all have relational implications (Fonagy et al., 2002). Most structured treatments include explicit psychoeducation and the use of case formulations, which may be important factors in the early development of alliance—establishing mutual agreement on aims and tasks in therapy. In psychotherapy processes, therapist empathy is a recognized facilitating factor, contributing to the bond between patient and therapist. MBT manuals emphasize that the patient needs to be validated and understood before being challenged on maladaptive patterns (Karterud et al., 2020; Karterud & Bateman, 2010), and the recommended therapeutic stance is to be mentalizing and curious as well as genuine and non-judgmental. Correspondingly, in Schema-Focused Therapy (SFT), mutual trust and positive regard (Bonds) are emphasized as important

alliance elements (Young et al., 2006), underlining the importance of an unthreatening, supportive therapist attitude. It is furthermore proposed that the SFT model itself promotes sympathy with the BPD patient (Young et al., 2006). A comparison study of SFT versus Transference-Focused Psychotherapy (TFP) indicated an increase in therapeutic alliance during both treatments (Spinhoven et al., 2007). In a study of Dialectical Behaviour Therapy (DBT), higher patient-rated therapy commitment and working capacity was associated with fewer suicide attempts (Bedics et al., 2015). As of yet, we have not found studies investigating relations between MBT alliance and outcomes.

The overriding aim of the current study was to investigate how aspects of therapeutic alliance (Goals, Bonds and Tasks) developed over time in MBT for patients with BPD. The study primarily aimed to investigate alliance processes in therapies with different clinical outcomes, and secondarily to explore variation associated with different patient characteristics.

Material and Methods

Design

The study is a quantitative, observational study with a longitudinal design.

Subjects

The studied sample included 155 BPD patients treated in an MBT unit during 2009–2016. Patients were referred on a regular basis to the outpatient clinic, which was on a specialist mental health service level, situated within a university hospital setting.

Mentalization-Based Treatment (MBT)

MBT was an outpatient treatment in accordance with MBT manuals (Karterud, 2011, 2012; Karterud & Bateman, 2010). The first year included weekly sessions of individual and group therapy and a psychoeducational group (12 sessions). Frequencies of individual therapy were gradually reduced in the second and third year, while group sessions continued throughout treatment. Treatment had an upper time limitation of 36 months.

Therapists

The team included three psychiatric nurses, three psychiatrists, an art therapist, a physiotherapist, a social worker, and two psychologists. Eight were qualified group analysts—one in psychoanalysis, one in

individual psychodynamic psychotherapy—67% were females, and mean age (year 2009) was 53 (*SD* = 9) years. Other individual therapists within the research period were different resident doctors and psychologists in training. All had basic MBT training and attended weekly video-based supervision by qualified MBT supervisors.

Therapist MBT Fidelity

MBT adherence and competence was assessed by video-recorded therapy sessions using the MBT Adherence and Competence Scale (Karterud et al., 2013) and the Adherence and Competence Scale for Mentalization-based Group Therapy (Folmo et al., 2017). On a 1–7 scale, a score of four or higher indicates adequate MBT adherence/competence. In 2013–2015, five raters evaluated 19 individual sessions (eight therapists) and 9 group sessions in the programme. For individual therapists, the mean adherence level was 4.7 (*SD* = 1.2) and the mean MBT competence level was 4.4 (*SD* = 1.2) (Kvarstein et al., 2019). For group therapists, the mean adherence level was 5.1 (*SD* = 1.37) and competence level 4.9 (*SD* = 1.30) (Kvarstein et al., 2020). This is comparable to a recent RCT study of MBT in groups for adolescents with BPD (Beck et al., 2020).

Baseline Assessment of Diagnoses

The MBT unit was part of a collaborative cross-regional network for treatment and research on personality disorders where all units used standardized measures for diagnostic assessment. Diagnoses were decided in accordance with the DSM-IV using the Mini International Neuropsychiatric Interview (MINI; Lecrubier et al., 1998) for symptom disorders and for PDs and the Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II; First et al., 1994). Assessments were performed by clinical staff with systematic training provided by the network in diagnostic interviews and principles of the LEAD-procedure (Longitudinal, Expert, All-Data; Pedersen et al., 2013). Evaluations were concluded with a psychiatrist.

Baseline Assessment of Other Patient and Treatment Factors

Self-reports of personality functioning, life quality, and work/study functioning (patient factors) and information about treatment termination (treatment factors) included the following:

Table I. Working alliance inventory—items according to WAI-S and WAI-SR.

WAI-S

Tasks	Goals	Bonds
1. __ and I have established a good understanding of what I need to do in treatment in order to improve my situation	4. __ does not understand what I am trying to accomplish in therapy	3. I believe___likes me.
2. What I am doing in therapy gives me new ways of looking at my problem.	6. __and I are working towards mutually agreed upon goals.	5. I trust that __ is able to help me.
8. __ and I have established a good understanding of what is important for me to work on.	10. __ and I have different understanding of my problems	7. I feel that___appreciates me.
12. I believe the way we are working with my problem is correct.	11. __and I have established a good understanding of the kind of changes that would be good for me.	9. __and I trust each other.

WAI-SR

Tasks	*Goals*	*Bonds*
1. As a result of these sessions I am clearer as to how I might be able to change.	4. __and I collaborate on setting goals for my therapy.	3. I believe___likes me.
2. What I am doing in therapy gives me new ways of looking at my problem.	6. __and I are working towards mutually agreed upon goals.	5.___and I respect each other.
10. I feel that the things I do in therapy will help me to accomplish the changes that I want.	8. __ and I agree on what is important for me to work on.	7. I feel that___appreciates me.
12. I believe the way we are working with my problem is correct.	11. __and I have established a good understanding of the kind of changes that would be good for me.	9. I feel _____ cares about me even when I do things that he/she does not approve of.

(1) Severity Indices of Personality Functioning —Short Form (SIPP-SF; Rossi et al., 2017) is a 60-item version of the original SIPP-118 (Pedersen, Arnevik, et al., 2017). It includes five personality problem domains. In this study the three domains with greatest impairment were: Identity Integration (12 items, aspects of enjoyment, meaning, self-esteem, and self-perception), Relational Capacity (10 items, aspects of attachment, intimacy, enjoying relationships, feeling appreciated, and being affectionate), and Self-control (12 items, aspects of controlling emotional reactions and impulsive behaviours). The remaining SIPP-SF domains are Responsibility and Social Concordance (Normal range T-scores: 40–60).

(2) EuroQol (EQ-5D) evaluates subjective life quality along five health dimensions and a global index (0–1). In the general population in Western societies, the global index score range is 0.80–0.89 (Saarni et al., 2007) and in PD populations 0.56 (Soeteman et al., 2008).

(3) The number of months they worked or studied at least 50% during the previous year was recorded to indicate current work functioning status by the patients.

(4) Reasons for treatment termination were recorded by the therapist.

Repeated Assessment of Working Alliance

The Working Alliance Inventory-Short (WAI-S; Tracey & Kokotovic, 1989), which is based on the Working Alliance Inventory (WAI; Horvath & Greenberg, 1989), was applied regularly after three months, then every sixth month during treatment, and once more at discharge. Due to the publication of the revised version (WAI-SR) by Hatcher and Gillaspy (2006), the WAI-S was replaced by WAI-SR in 2012. Both WAI versions include three subscales— Goals, Bonds, and Tasks—with four items for each. Each item is rated on a 7-point scale, and scores above 4 signify satisfactory alliance. Table I presents differences between WAI versions. In the following, we use the term WAI-S/SR. The possible impacts of differences in WAI versions were investigated using a categorical variable identifying subgroups with only the WAI-S version, only the WAI-SR version, and both versions (longitudinal data series with both WAI-S and WAI-SR). In the first part of the study, patients received the WAI-S ($n = 34$). From June 2012, admitted patients had only WAI-SR measures ($n = 71$). The remaining patients had longitudinal data-series including both versions.

Assessment of Clinical Outcomes

(1) The observer-rated Global Assessment of Functioning (GAF) provides a composite score combining social and symptom-related

impairment and was therefore chosen as a global outcome measure (0–100 scale, Axis V, DSM-IV) (Pedersen et al., 2018). GAF evaluations were performed by staff therapists (baseline, after three months, repeatedly every sixth month throughout treatment, and at treatment termination). All therapists working at the unit received a systematic GAF training course. The validity and generalizability of GAF scores was previously investigated within several treatment units representing the same clinical contexts and included therapists at the specific unit (Pedersen et al., 2007). The study demonstrated high consistency of GAF scores across different raters and also different treatment units (generalizability coefficients of absolute decision (the score) range .86 to .95).

High GAF scores indicate better psychosocial functioning; scores above 60 represent mild/no impairment (Pedersen et al., 2018). In this study, the sample ($n = 155$) was divided into two subgroups according to outcomes: (i) the poor outcome subgroup (GAF below 60 at the end of treatment; 43%) and (ii) the good outcome subgroup (GAF equal or above 60 at the end of treatment; 57%). All 155 patients had a baseline GAF assessment, and 7 patients lacked a final GAF assessment on treatment termination.

(2) In order to supplement GAF as an observer-rated instrument, outcomes additionally included two patient-reported measures also administered at baseline, after three months, repeatedly every sixth month throughout treatment, and at treatment termination. All self-report measures (outcomes and alliance) were administered by the secretary at the unit. The profiles from self-reports constituted a basis for clinical evaluation of treatment progress.

(a) The Work and Social Adjustment Scale (WSAS) is a validated 5-item self-report measure of functional impairment (i.e., work, social and private leisure activities, domestic work, and close relations) (Pedersen, Kvarstein, et al., 2017). Total sum-scores below 15 represent mild/no impairment (Mundt et al., 2002); these scores were found among 58% of the patients who filled in self-reports at the end of treatment. Among these, 85% were also in the good outcome GAF subgroup. All 155 patients had the baseline assessment, but 42 patients lacked a final WSAS assessment on treatment termination.

(b) The BSI-18 is a self-report measure derived from the 53-item Brief Symptom Inventory (BSI), a shortened form of the Symptom Checklist-90-Revised (SCL-90-R) (Derogatis, 2000). The BSI includes 18 items and assesses symptom distress (depression, somatization, anxiety) on a 0–4 scale. Non-clinical distress is indicated by a mean sum-score of 0.8 (Pedersen & Karterud, 2004), which was found among 48% of the patients who filled in self-reports at the end of treatment. Among these, 78% were also in the good outcome GAF subgroup. All 155 patients had the baseline assessment, but 42 patients lacked a final BSI assessment on treatment termination.

The process and outcome measures used in this study were a part of the standardized assessment and treatment evaluation used within the collaborative cross-regional network for treatment and research on personality disorders.

Ethics

All patients gave written, informed consent to participate in the research. The treatment unit collected clinical data, which was registered in an anonymous database administrated by Oslo University Hospital. Procedures for data collection ensured that participating individuals could not be identified. Data security systems were approved by the Data Protection Official at Oslo University Hospital. Based on anonymous data, ethical approval was not required from the Regional Committee for Medical Research and Ethics.

Statistical Procedures

Hierarchical models (mixed models) (Singer & Willett, 2003) were used for statistical analyses of longitudinal data (mixed models, IBM SPSS statistics version 25) in order to maximize utilization of available patient data and capture change over time. Time (months from baseline) was modelled as a continuous variable in all models. Linear trajectories captured significant longitudinal trends for all dependent variables, among which WAI-S/SR was the main dependent variable, and GAF, WSAS, and BSI represented preliminary analyses ($p < 0.001$). Log likelihood estimations of model fit indicated significant improvements from an unconditional model to a linear random coefficients (intercept and slope) model (critical values for chi-square statistic: $p < 0.01$) using an unstructured covariance type. The equation was: $Y_{ij} = \beta_0 + \beta_1 \, time_{ij} + b_0 + b_1 \, time_{ij} + \varepsilon_{ij}$.

Y is the dependent variable for all observations, individuals (i), and assessment times (j), β is the fixed effects regression coefficient, b the random effects regression coefficient, and ε indicates residual variation. The categorical variable for different WAI versions was investigated in the three WAI-S/SR subscale models; some reduction of longitudinal variation (3, 7, and 5%, respectively) was discovered, but estimate deviations were not statistically significant ($p > 0.05$) and did not improve AIC estimations of model fit.

Initial analyses investigated the longitudinal change of clinical outcome variables and variation

Table II. Baseline characteristics for MBT patients in subgroups with good and poor outcomes.

	Good outcome		Poor outcome	
	Mean (SD)	%	Mean (SD)	%
Personality disorder				
Borderline PD		73		73
Borderline PD traits	5.5(1.9)		5.4(1.9)	
Total number of PD traits	14.0(5.8)		14.5(5.8)	
Number of PDs	1.3(0.6)		1.4(0.7)	
Other PDs than BPD:				
Paranoid PD		7		16*
Narcissistic PD		4		3
Antisocial PD		0		5
Avoidant PD		18		19
Obsessive Compulsive PD		8		9
Dependent PD		2		6
NOSPD		15		17
Severity Indices Personality Problems (T-scores)				
Self control	23(12)		25(13)	
Social concordance	32(13)		32(15)	
Identity	20(10)		22(10)	
Relation	30(10)		31(12)	
Responsibility	27(14)		29(13)	
Comorbid symptom disorders				
Total number, symptom disorders	2.4(1.4)		2.7(1.5)	
Mood		65		83*
Anxiety		74		71
PTSD		17		18
OCD		4		7
ADHD		1		2
Eating		19		22
Substance abuse		24		14
Baseline status				
Age	26(6)		30(6)*	
Gender female		85		81
No work/study at all last year		27		51*
Global Index Life Quality EQ 5D-3L	46(20)		42(18)	

Notes: Descriptive data with mean values, standard deviations (SD), and valid per cent (%). Poor outcomes were defined as GAF end-score < 60 ($n = 64$, total $N = 155$). Significant differences are marked with * ($p < 0.05$, Pearson chi-square test/independent samples T-test).

associated with selected patient and treatment factors as predictors. The investigated patient and treatment factors were chosen on the basis of descriptive sample characteristics in the two outcome subgroups (Table II) and are given in Table III. The main investigation included longitudinal change of WAI-S/SR subscales and variation associated with the dichotomous GAF variable as a predictor, as well as a moderator interaction combining the dichotomous GAF variable and patient factors. The equations for the predictor analyses were: $Y_{ij} = \beta_0 + \beta_1 \text{time}_{ij} + \beta_2 \text{PRED} + \beta_3 \text{PRED time}_{ij} + b_0 + b_1 \text{time}_{ij} + \varepsilon_{ij}$. The equations for the moderator analyses were: $Y_{ij} = \beta_0 + \beta_1 \text{time}_{ij} + \beta_4 \text{PREDMOD} + \beta_5 \text{PREDMOD time}_{ij} + b_0 + b_1 \text{time}_{ij} + \varepsilon_{ij}$. For interpretation of predictors/interactions; each model is judged by the predictor-associated deviation of the trajectory for the dependent variable (estimated deviation of intercept level and slope/change-rate), explained variation (% change from the estimated variation in the initial linear random coefficients model), and change in estimates of log likelihood statistics (indices of model fit, Akaikes Information Criterion, AIC).

The sample had unbalanced data with different numbers of assessments per patient. The chosen method for longitudinal analyses incorporates unbalanced data and uses all available data for each individual trajectory. Therefore, in this study, we did not use imputation to compensate for missing data. All included patients had at least one assessment. The mean number of WAI-S/SR assessments was 3.2 ($SD = 1.8$, range 1–9), 48% lacked the final assessment, 22% had only one assessment, 58% had at least three, and 38% had four or more during the course of treatment. Among the 34 patients with only one assessment, 53% were early alliance assessments performed within the first year. A variable counting numbers of assessment points for each individual captured a relevant missing data pattern. To investigate the effect of this missing data pattern on the outcomes, the variable was added as a predictor in all three working alliance subscale models (Hedeker & Gibbons, 1997). Analyses indicated poorer initial alliance ratings for patients with fewer assessments (for all working alliance subscales $p < 0.05$), but no significant effect of the variable on alliance development over time ($p > 0.05$, all subscales). As longitudinal effects were of primary interest, we did no further overall estimation of missing-data effects.

All final longitudinal working alliance models included (a) investigations controlling for differences in WAI versions, (b) investigations controlling for different numbers of assessment, and (c) models

Table III. Linear mixed model estimations: Clinical outcomes and variation for patients in MBT.

LMM		Intercept estimate	Linear slope estimate	Explained intercept variation	Explained slope variation	Model fit
Model	Predictor	Mean (SE)	Mean (SE)	%	%	AIC
WSAS		25.3(0.6)	−0.3(0.03)			3977
BSI		2.08(0.06)	−0.03(0.003)			1590
GAF		49(0.4)	0.4(0.03)	*Reference**	*Reference**	4532
	Patient factors					
	Age	ns	−0.01(0.004)*	0	5	4531
	Work/study mnths last yr	ns	ns	0	0	4171
	Paranoid PD	ns	ns	1	2	4533
	Mood disorder	2.02(0.9)	ns	9	0	4502
	Treatment factors					
	Completed as planned	ns	−0.3(0.06)***	2	25	4509

Notes: Linear mixed model estimations with GAF, WSAS, and BSI as dependent variables. Intercept and slope estimates are given for each model. The variance estimates in each model are the reference values for calculating explained variance for each investigated predictor. Indicator of model fit is Akaikes Information Criterion (AIC). The table presents estimated deviance of intercept and slope and explained variance associated with patient and treatment factors in each model. Significant differences are marked with * ($p < 0.05$) **($p < 0.01$) or ***($p < 0.001$). A significant variation estimate in the initial model ($p<0.05$) is given by *Reference* *.

investigating the supplementary dichotomous WSAS and BSI outcome variables.

Results

Descriptive Data in Subgroups with Different Outcomes

Patient factors. The vast majority had BPD with severe disorder, indicated by poor life-quality, considerable comorbidity, and personality problems across all domains, although most severe within the domains of identity and self-control (Table II). Differences by GAF outcome subgroup were minor ($p > 0.1$, independent samples T-test) with the exception of younger age, fewer patients with no months of work/study at all previous year, and fewer with comorbid Paranoid PD and mood disorder in the good outcome subgroup ($p < 0.05$, independent samples T-test).

Treatment factors. Mean treatment duration was 27 months (SD 13), early drop out (<6-month duration) was minimal (2.5%), and neither differed by subgroup ($p > 0.05$, independent samples T-test/ Pearson chi square test). Nearly all patients in the good outcome subgroup completed treatment according to plan (91%) versus 58% in the poor outcome subgroup ($p < 0.05$, Pearson chi square test). In the good outcome group, there were no later drop-outs, termination was advised for 2%, and no patients were referred to other institutions. In the poor outcome group, 9% were later drop-outs, termination was advised for 12%, and 13% were referred to other institutions when they terminated treatment.

Clinical Outcomes and Factors Associated with Longitudinal Outcome Variation

Baseline levels indicated severe problems of functioning and distress at the start of MBT, and significant variation was found within the longitudinal data for all three outcome variables. Overall, corresponding and significant longitudinal improvement over time was found for GAF, WSAS, and BSI (Table III). Mean change was for GAF$_{admission}$ 48.1 (SD 5.5) to GAF$_{discharge}$ 61.7 (SD11.7) ($n = 148$), for WSAS$_{admission}$ 25.9 (SD 7.9) to WSAS$_{discharge}$ 13.7 (SD 10.7) ($n = 107$), and for BSI$_{admission}$ 2.1 (SD 0.8) to BSI$_{discharge}$ 1.2 (SD 0.9) ($n = 107$). Among significant treatment factors (Table III), completing treatment according to plan was strongly associated with better GAF improvement, and this explains 25% of the GAF slope variation (Table III). Age and paranoid PD were also noteworthy patient factors, explaining 2–5% of the GAF slope variation. Higher age was associated with significantly poorer GAF improvement (Table III). Mood disorder was significantly associated with baseline GAF but did not explain further variation of GAF development over time (Table III). These preliminary analyses suggested that intrinsic treatment factors were relevant for further investigation and also pointed to certain patient factors. We thus proceeded with investigations of the main dependent variable, WAI-S/SR, as a potential indicator of intrinsic treatment quality,

with the dichotomous GAF subgroup variable as an indicator of outcome variation.

Main Analyses: Longitudinal Course of Working Alliance

Overall, patients rated high initial levels of working alliance. Goals, Bonds, and Tasks all had initial levels well within a satisfactory range, and the overall picture of change over time in the sample was a significant increase of all three working alliance subscales (Table IV). Analyses also revealed significant longitudinal between-subject variation. These change patterns also remained significant in models (a) controlling for variation associated with different WAI versions and (b) investigating possible bias of different assessment numbers.

Variation associated with good and poor outcome subgroups. The good and poor outcome subgroup predictor was investigated in each of the three models. Initial levels of Goals, Bonds, and Tasks did not differ by subgroup, but change over time was significantly different by subgroup (Table IV; Figure 1). The subscale Goals accounted for 23% of the WAI-S/SR slope variation, Bonds for 25%, and Tasks for 35% (Table IV). Results remained significant ($p < 0.05$; all analyses used linear mixed models) for the three subscales—Goals, Bonds, and Tasks—in models (a) controlling for variation associated with different WAI versions, (b) investigating possible bias of different assessment numbers, and (c) corresponding differences were also found in models investigating the dichotomous WSAS and BSI outcome variables as predictors. Further investigation dividing the sample by outcome subgroup revealed the following. Including only patients in the good outcome subgroup, ratings of Goals, Bonds, and Tasks increased significantly over time (for all $p < 0.05$). Including only patients in the poor outcome subgroup, change over time was not significant for any of the WAI-S/SR subscales (for all $p > 0.1$).

Variation associated with patient factors. Relevant patient factors (age, comorbid mood disorder, and comorbid paranoid PD) were investigated as separate predictors added to the three WAI-S/SR subscale models. Age was not associated with significantly deviating initial alliance levels or deviating change over time (all subscales, $p > 0.1$), but explained some longitudinal variation. Paranoid PD was not significantly associated with initial alliance levels (all subscales, $p > 0.1$), but associated with significantly less improvement of two of the three

subscales over time (Tasks and Bonds, $p < 0.05$). Although the slope deviation of Goals was not significant, paranoid PD explained 8% of the slope variation of this subscale ($p = 0.11$). Mood disorder was associated with significantly lower initial alliance levels, but not deviating change over time (all subscales, $p > 0.1$). Table IV demonstrates estimates for the subscale Tasks, also illustrated in Figure 2.

Further investigation included the moderator interaction between outcome subgroups and (i) paranoid PD and (ii) age. Paranoid PD was not associated with baseline deviation of WAI-S/SR ratings in any of the two outcome subgroups ($p > 0.1$, all three subscales). The presence of paranoid PD was associated with significantly poorer development of WAI-S/SR subscales over time in the poor outcome subgroup ($p < 0.05$, all subscales), but not in the good outcome subgroup ($p > 0.1$, all subscales). The impact of age on alliance development was not further explained by differentiation according to outcome subgroup ($p > 0.1$, all subscales).

Significant results remained ($p < 0.05$) in models controlling for variation associated with different WAI versions. In models investigating possible bias of different assessment numbers, the trend of poorer development of WAI-S/SR subscales over time was less prominent ($p < 0.1$, all subscales). Corresponding results for paranoid PD were also found in models investigating the supplementary dichotomous WSAS and BSI outcome variables.

Discussion

There is little research on alliance and outcomes in specialized treatments for BPD. This study represents a large sample of patients attending a mentalization-based treatment (MBT) programme in an outpatient format in a regular, not an experimental, treatment setting. It is among the first longitudinal studies of alliance in MBT and captures patterns of early alliance in a large sample of patients with BPD.

Main Findings

Overall, patient-reported working alliance in MBT was initially within a satisfactory range (scores above 4). During therapy, all subscales—Goals, Bonds, and Tasks—increased over time. Differentiation in subgroups with good and poorer outcomes revealed the following:

(1) Initial working alliance—Goals, Bonds, and Tasks—did not differ by outcome.

Table IV. Working alliance in subgroups with high or poor end GAF in MBT individual therapy.

Model	Predictor	Moderator	Intercept	Linear Slope	Explained intercept variation	Explained slope variation	Model fit
			Mean (SE)	Mean (SE)	%	%	AIC
GOALS			4.98(0.1)	0.02(0.01)	*Reference**	*Reference**	1449
	Poor vs high outcome		ns	−0.02(0.01)*	0	23	1434
BONDS			5.14(0.1)	0.02(0.004)	*Reference**	*Reference**	1453
	Poor vs high outcome		ns	−0.02(0.01)*	8	25	1432
TASKS			5.04(0.11)	0.02(0.01)	*Reference**	*Reference**	1496
	Poor vs high outcome		ns	−0.03 (0.01)***	8	35	1470
	Age		ns	ns	0	5	1496
	Mood disorder		−0.54 (0.3)*	ns	8	0	1484
	Paranoid PD		ns	−0.03(0.01)*	8	10	1496
	Poor outcome	Paranoid PD=1	ns	−0.05(0.02)*	8	45	1470
	Poor outcome	Paranoid PD=0	ns	ns			
	High outcome	Paranoid PD=1	ns	ns			
	High outcome	Paranoid PD=0	ns	ns			

Notes: Linear mixed model estimations with WAI-S/SR subscales as dependent variables. Intercept and slope estimates are given for each model. The variance estimates in each model are the reference values for calculating explained variance for each investigated predictor. Indicator of model fit is Akaikes Information Criterion (AIC). The table presents estimated deviance of intercept and slope and explained variance associated with outcome subgroups in each model, and estimated deviance of intercept and slope and explained variance associated with patient factors for the model with the subscale TASKS as dependent variable. Significant differences are marked with * ($p < 0.05$) ** ($p < 0.01$) or *** ($p < 0.001$). A significant variation estimate in the initial model ($p < 0.05$) is given by *Reference **.

(2)   Positive temporal development of therapeutic alliance during therapy characterized good outcomes.

(3)   Outcome subgroups differed most in the development of the Tasks subscale.

(4)   Comorbid paranoid PD was more frequent in the subgroup with poor outcomes and associated with poorer alliance development in this subgroup.

In Bordin's (1979) operationalization of alliance, an important formative and collaborative aspect of the process takes place in the initial phase—the agreement upon Goals in therapy. In MBT, treatment Goals are defined in an early case formulation tailored for the individual patient. In line with other psychotherapy research, MBT also emphasizes the importance of patients' own understanding in the negotiation of work in therapy (Muran & Barber, 2011). As poorly functioning patients may have difficulties formulating or understanding the concepts of therapy, MBT recommends clear, simple, and short formulations. Given the emphasis of relational problems among BPD patients, the early alliance ratings in this study were high—perhaps more so than could be expected. However, others have also demonstrated high alliance

levels in psychotherapy despite severe interpersonal problems (Ollila et al., 2016). The authors discuss how patients with extensive interpersonal problems might feel a strong need for help and be motivated to engage in the process. In the present sample, initial alliance levels may likewise reflect positive expectations in the start of therapy.

Being referred to an extensive MBT programme is often preceded by several former treatment attempts. Illustrating this point, a recent qualitative study of MBT emphasized how patients experienced a positive shift of expectations when starting to engage in therapy (Gardner et al., 2020). Expectations were nevertheless ambiguous, ranging from seeing MBT as potentially life-saving to perceptions of MBT as "a last chance saloon." The latter illustrates how the early alliance may also be extremely fragile. In the current poorly functioning BPD sample, self-reported relational problems, covering issues of attachment and intimacy, were notable. Characteristically, relational problems among BPD patients represent a strong need for close relationships together with high interpersonal sensitivity and overwhelming fear of rejection.

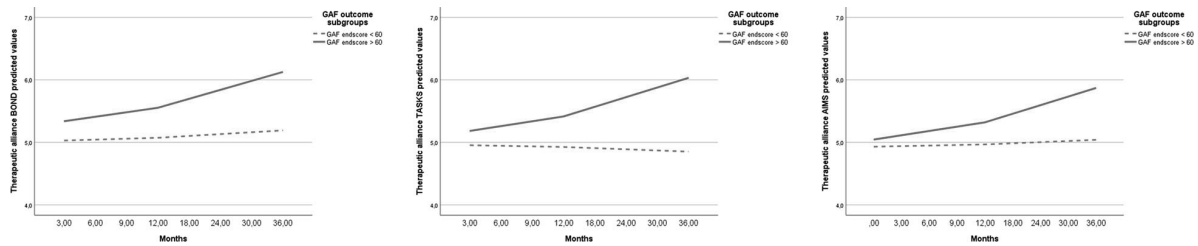In the present study, poorer alliance levels were related to comorbid mood disorder, and overall,

Figure 1. The alliance development in subgroups with different outcomes.

patients' personality problems reflected low levels of self-esteem, enjoyment, and meaning. Comorbidity of BPD and mood disorder is generally common—the two conditions share vulnerabilities and may be hard to distinguish (Skodol et al., 2010). This study demonstrates that the collaborative starting point of therapy was negatively influenced by dysphoria. Interestingly, over time, its negative effect was not enduring. Although viewed as a potential setback at onset, comorbid mood disorder did not complicate a treatment process focusing on core personality problems.

Moreover, the subgroups with different clinical outcomes did not differ with respect to their initial ratings of alliance. Attachment processes in treatment are assumed to take time. The present study captures the development of alliance among BPD patients in a specialized treatment—MBT. The results confirm a positive development for the majority but also that the development of alliance over time is a vulnerable process. In treatments of BPD, the capacity to develop a working alliance suggests a process of

gaining epistemic trust and relational competence. As a whole, a positive development of the working alliance was indeed a main trend in the present sample, as were favourable clinical improvements. Correspondingly, positive clinical outcomes were demonstrated in a former study of a smaller MBT sample within the same treatment context but reflecting a shorter study-period (Kvarstein et al., 2015). This study did not include measures of alliance, but low early drop-out rates may, nevertheless, be indicative of satisfactory initial bonding.

The MBT manual instructs the therapist to be an attachment figure (Karterud & Bateman, 2010), and as BPD patients display substantial attachment issues, these will be central in the further development of a working alliance. The MBT manual emphasizes that this process requires emotional involvement from the therapist (Karterud & Bateman, 2010). Self-perception and self-esteem are both aspects of mentalizing that may be developed in a validating and reflective attachment context. Irrespective of outcome, patients in our study reported
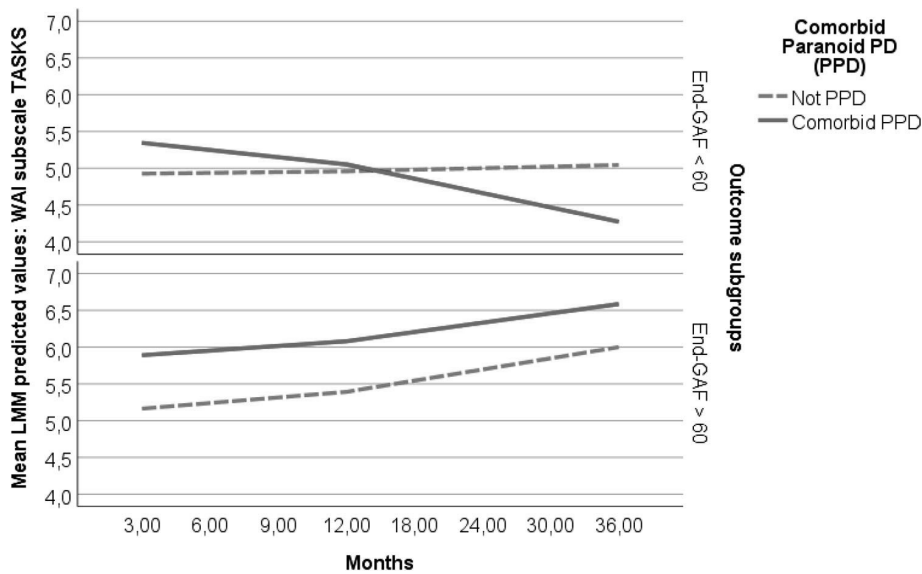

Figure 2. Comorbid paranoid PD and working alliance in subgroups with different outcomes.

severe personality problems of both self-aspects. In the present study, the majority of patients achieved a positive process with clinical improvement together with a growing working alliance. However, within this sample, less favourable processes were also evident.

The present results highlight how patients need to understand what psychotherapy is and see how it may work. Differences between outcome subgroups were particularly marked for the subscale Tasks. The specific items indicate confidence in the treatment method, clarity on the therapeutic work being done, and a sense of already noticing personal change. As Bordin (1979) stated, alliance is the ingredient that " … makes it possible for the patient to accept and follow treatment faithfully" (p. 2). In our study, initial experiences of tasks in MBT did not differ by subgroup, but emerged over time—for most patients, the process was fruitful.

A reasonable question is to what extent poorly mentalizing patients are able to understand how to work in therapy in the best way (Dimaggio et al., 2019). This is the main argument for the systematic use of both case formulations and psychoeducation in specialized BPD treatments. Both aim to increase patients' knowledge about BPD, such as different problems of mentalizing, recognition and regulation of affects, or understanding patterns of relational attachment. The psychoeducation also introduces the treatment programme, what is expected of patients and therapists, and what kinds of things therapists may ask about (Karterud, 2011). In qualitative interviews of psychoeducation in MBT, patients reported the importance of feeling understood, often for the first time, and most essentially, experiencing hope for change (Ditlefsen et al., 2020). The study also points to negative experiences of feeling too different from other patients in the group. Positive experiences of validation and learning could indicate a strengthening of epistemic trust (Fonagy et al., 2015). Mutual agreement on problems and provision of the rationale behind a treatment method is considered crucial for the outcome (Wampold & Imel, 2015). In treatment of BPD, emotional dysregulation and high-risk impulsivity are often part of a challenging picture. In the present study, alliance developed despite substantial initial problems of self-control.

For patients with positive developments, it seems they became able to collaborate (Goals and Tasks), and cocreate a trusting relationship (Bonds). A qualitative study of therapist interventions in MBT described how therapists within a good working alliance context not only validated and supported the patient but could also challenge maladaptive patterns (Folmo et al., 2019). A positive alliance process could indicate that the patient gradually comes to understand the importance of working on the problems focused in therapy and becomes increasingly willing to be challenged directly on these matters. Therapy then becomes increasingly relevant—permitting work on central personality problems. In a study comparing alliance for BPD patients in SFT and TFP (Spinhoven et al., 2007), method-specific factors influenced the quality of the alliance. SFT, with its emphasis on the "necessary and sufficient conditions" in the client-centred approach, produced the better alliance, whereas the first stages of TFP, in which aggressive self- and object representations are activated and interpreted, demanded too much of the early alliance.

Results indicate that the long-term therapy process could also be cumbersome. The capacity for gaining mutual understanding is essential in therapy dyads, group treatments, and human interaction in general. Paranoid PD was characterized by a limited collaborative alliance process. Few have investigated the effect of MBT on other personality disorders (PD) besides BPD (Volkert et al., 2019). However, the clinical severity of PD in terms of comorbidity has been investigated. Studies of social cognition among patients with BPD have indicated that more impaired mentalizing is associated with more severe, comorbid PD (Normann-Eide et al., 2019), and MBT studies differentiating between patients with only BPD and patients with PD comorbidity recommend MBT for the more complex conditions (Kvarstein et al., 2019). BPD with comorbid paranoid PD represents a common form of severe personality pathology and can be conceptualized within the frame of epistemic trust. Counterintuitively, our study did not demonstrate differences in initial alliance or GAF levels related to comorbid paranoid PD. However, paranoid PD was overrepresented in the subgroup with poorer clinical improvement and associated with impeded alliance development over time; it is quite possible that many patient-therapist dyads were unable to handle alliance ruptures adequately. However, the present study also signals the possibility of a positive course. Paranoid PD was not associated with impeded alliance in the good outcome group.

In treatment of severe PD, the alliance process depends on the quality of the dyad between the patient—who, in the case of paranoid features, may be reserved, hostile, or dismissing—and the therapist—who, in such cases, has to keep up engagement and manage countertransference activated by rejection, criticism, or devaluation. It is plausible that adhering to a specific treatment model and strategy, such as MBT, could provide a helpful framework. As advanced in MBT, a genuine and frank style of communication may prevent paranoid phantasies

about the therapist's thoughts or intentions. An empathic, dyadic process of enquiry and reflection may also build confidence. A mentalizing process implies that therapists are reasonably transparent about their own mental states and that patients work on their understanding of self and others. Talented therapists may be more responsive and attentive to ways of facilitating alliance with their patients (Lemma et al., 2011; Wampold & Imel, 2015). Two studies of psychotherapy alliance recommended a combination of self-doubt as a therapist and self-affiliation as a person (Heinonen & Nissen-Lie, 2020; Nissen-Lie et al., 2017). In line with such findings, we may speculate that being willing to be transparent (e.g., display self-doubt) whilst having sufficient self-affiliation not to be overwhelmed by countertransferences (feeling devalued, mistrusted, rejected, etc.) is crucial for fostering alliance with patients presenting paranoid PD.

There are many possible pitfalls in such processes. Studies of psychotherapy with relationally disturbed patients have indicated high sensitivity towards therapists' countertransference reactions or behaviours. Negative therapist feelings of disengagement or inadequacy are associated with poorer outcomes, and therapists' anxiety or negative reactions may contribute to a poorer working alliance (Dahl et al., 2016, 2017; Nissen-Lie et al., 2015). Moreover, too great a degree of self-confidence or self-concern may not facilitate alliance (Heinonen & Nissen-Lie, 2020; Nissen-Lie et al., 2010). In order to support therapist competence, MBT recommends transparency, not only within sessions, but including active collaboration with a team of therapists and regular supervision—all to ensure a mentalizing culture on all levels. It is noteworthy that studies have demonstrated that outcomes in MBT for poorly functioning patients depend on overall quality of both therapist competence and treatment organization (Bales, Timman, et al., 2017; Bales, Verheul, et al., 2017).

Strengths and Limitations

The sample represented a large, clinically representative, and severe BPD population of 155 patients treated in an MBT programme, and the study has a longitudinal design. Few MBT studies include fidelity measures and as such are often neglected in psychotherapy research (Perepletchikova et al., 2007). As is often the case in clinical studies, longitudinal data were unbalanced with different numbers of assessments. Assessments were performed during treatment, and one reason for different assessment numbers is different treatment durations. To compensate, assessments at termination were placed at the last 36-month time-point. We chose a maximum likelihood-based statistical method for longitudinal analyses, generating individual curves based on all available data for each patient. Different assessment numbers were not associated with longitudinal deviation, and all reported results remained evident when we investigated possible bias of different assessment numbers. The study included two different versions of WAI, and differently formulated items may have affected the alliance ratings. We present both versions and include investigations controlling for possible impacts of WAI. The limitation is considered minor, as we found little conceptual difference between the two WAI versions. Different WAI versions were not associated with significant longitudinal deviation, and all reported results remained evident when we controlled for different WAI versions. As a study of alliance, it is limited in that it only includes patient ratings. Even though patient- and therapist-rated alliance are equally good predictors of outcome (Flückiger et al., 2018), our findings are restricted by the fact that we only study patient-rated alliance. There could be a tendency to both over- and underestimate the alliance based on self-report only (Tryon et al., 2008), thus potentially making our results less reliable. Moreover, more frequent measures of alliance would have been a better basis for longitudinal trends and would have compensated for the possible bias of session to session fluctuations. However, the study design was pragmatic, as it investigated an ongoing treatment, and research was based on assessments, which were part of the unit's regular clinical evaluation routines.

Conclusion

The study demonstrates satisfactory levels of initial working alliance among BPD patients in MBT irrespective of clinical outcomes and an overall increase of all alliance aspects over time. Further investigation revealed that comorbid paranoid PD was more frequent in the subgroup with poor outcomes and associated with poorer alliance development in this subgroup. Differences in alliance development according to outcome were most pronounced for the subscale tasks.

ORCID

E. J. Folmo http://orcid.org/0000-0001-8608-3780
M. S. Johansen http://orcid.org/0000-0001-8608-3780

References

Bales, D. L., Timman, R., Luyten, P., Busschbach, J., Verheul, R., & Hutsebaut, J. (2017). Implementation of evidence-based treatments for borderline personality disorder: The impact of organizational changes on treatment outcome of mentalization-based treatment. *Personality and Mental Health*, *11*(4), 266–277. https://doi.org/10.1002/pmh.1381

Bales, D. L., Verheul, R., & Hutsebaut, J. (2017). Barriers and facilitators to the implementation of mentalization-based treatment (MBT) for borderline personality disorder. *Personality and Mental Health*, *11*(2), 118–131. https://doi.org/10.1002/pmh.1368

Bateman, A., & Fonagy, P. (2016). *Mentalization-based treatment for personality disorders: A practical guide*. Oxford University Press.

Beck, E., Bo, S., Jørgensen, M. S., Gondan, M., Poulsen, S., Storebø, O. J., Andersen, C. F., Folmo, E., Sharp, C., Pedersen, J., & Simonsen, E. (2020). Mentalization-based treatment in groups for adolescents with borderline personality disorder: A randomized controlled trial. *Journal of Child Psychology and Psychiatry*, *61*(5), 594–604. https://doi.org/10.1111/jcpp.13152

Bedics, J. D., Atkins, D. C., Harned, M. S., & Linehan, M. M. (2015). The therapeutic alliance as a predictor of outcome in dialectical behavior therapy versus nonbehavioral psychotherapy by experts for borderline personality disorder. *Psychotherapy*, *52*(1), 67–77. https://doi.org/10.1037/a0038457

Bo, S., Sharp, C., Fonagy, P., & Kongerslev, M. (2017). Hypermentalizing, attachment, and epistemic trust in adolescent BPD: Clinical illustrations. *Personality Disorders: Theory, Research, and Treatment*, *8*(2), 172. https://doi.org/10.1037/per0000161

Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research & Practice*, *16*(3), 252. https://doi.org/10.1037/h0085885

Colli, A., Tanzilli, A., Dimaggio, G., & Lingiardi, V. (2014). Patient personality and therapist response: An empirical investigation. *American Journal of Psychiatry*, *171*(1), 102–108. https://doi.org/10.1176/appi.ajp.2013.13020224

Dahl, H. S. J., Høglend, P., Ulberg, R., Amlo, S., Gabbard, G. O., Perry, J. C., & Christoph, P. C. (2017). Does therapists' disengaged feelings influence the effect of transference work? A study on countertransference. *Clinical Psychology & Psychotherapy*, *24*(2), 462–474. https://doi.org/10.1002/cpp.2015

Dahl, H.-S. J., Ulberg, R., Friis, S., Perry, J. C., & Høglend, P. A. (2016). Therapists' inadequate feelings and long-term effect of transference work. *Psychotherapy and Psychosomatics*, *85*(5), 309–310. https://doi.org/10.1159/000444647

Derogatis, L. (2000). *The brief symptom inventory-18 (BSI-18): Administration. Scoring, and procedures manual* (3rd ed.). National Computer Systems.

Dimaggio, G., Maillard, P., MacBeth, A., & Kramer, U. (2019). Effects of therapeutic alliance and metacognition on outcome in a brief psychological treatment for borderline personality disorder. *Psychiatry*, *82*(2), 143–157. https://doi.org/10.1080/00332747.2019.1610295

Ditlefsen, I. T., Nissen-Lie, H., Andenæs, A., Normann-Eide, E., Johansen, M. S., & Kvarstein, E. H. (2020). *"Yes, there is actually hope!"– A qualitative investigation of how patients experience mentalization-based psychoeducation tailored for borderline personality disorder*. Manuscript submitted for publication.

Dyson, H., & Brown, D. (2016). The experience of mentalization-based treatment: An interpretative phenomenological study. *Issues in Mental Health Nursing*, *37*(8), 586–595. https://doi.org/10.3109/01612840.2016.1155246

Euler, S., Wrege, J., Busmann, M., Lindenmeyer, H. J., Sollberger, D., Lang, U. E., Gaab, J., & Walter, M. (2018). Exclusion-proneness in borderline personality disorder inpatients impairs alliance in mentalization-based group therapy. *Frontiers in Psychology*, *9*, 824. https://doi.org/10.3389/fpsyg.2018.00824

Falkenström, F., & Larsson, M. H. (2017). The working alliance: From global outcome prediction to micro-analyses of within-session fluctuations. *Psychoanalytic Inquiry*, *37*(3), 167–178. https://doi.org/10.1080/07351690.2017.1285186

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. (1994). *Structured clinical interview for Axis I DSM-IV disorders*. Biometrics Research.

Flückiger, C., Del Re, A., Wampold, B. E., & Horvath, A. O. (2018). The alliance in adult psychotherapy: A meta-analytic synthesis. *Psychotherapy*, *55*(4), 316. https://doi.org/10.1037/pst0000172

Folmo, E. J., Karterud, S. W., Bremer, K., Walther, K. L., Kvarstein, E. H., & Pedersen, G. A. (2017). The design of the MBT-G adherence and quality scale. *Scandinavian Journal of Psychology*, *58*(4), 341–349. https://doi.org/10.1111/sjop.12375

Folmo, E., Karterud, S., Kongerslev, M., Kvarstein, E., & Stänicke, E. (2019). Battles of the comfort zone: Modelling therapeutic strategy, alliance, and epistemic trust—a qualitative study of mentalization-based therapy for borderline personality disorder. *Journal of Contemporary Psychotherapy*, *49*(3), 141–151. https://doi.org/10.1007/s10879-018-09414-3

Fonagy, P., Gergely, G., Target, M., & Jurist, E. L. (2002). *Affect regulation, mentalization, and the development of the self*. Other Press, LLC.

Fonagy, P., Luyten, P., & Bateman, A. (2015). Translation: Mentalizing as treatment target in borderline personality disorder. *Personality Disorders: Theory, Research, and Treatment*, *6*(4), 380–392. https://doi.org/10.1037/per0000113

Gardner, K. J., Wright, K. M., Elliott, A., Graham, S., & Fonagy, P. (2020). The weirdness of having a bunch of other minds like yours in the room: The lived experiences of mentalization-based therapy for borderline personality disorder. *Psychology and Psychotherapy: Theory, Research and Practice*, *93*(3), 572–586. https://doi.org/10.1111/papt.12243

Harpøth, T. S., Kongerslev, M. T., Moeyaert, M., Bo, S., Bateman, A. W., & Simonsen, E. (2019). Evaluating "mentalizing positive affect" as an intervention for enhancing positive affectivity in borderline personality disorder using a single-case multiple-baseline design. *Psychotherapy*. https://doi.org/10.1037/pst0000251

Hatcher, R. L., & Gillaspy, J. A. (2006). Development and validation of a revised short version of the working alliance inventory. *Psychotherapy Research*, *16*(1), 12–25. https://doi.org/10.1080/10503300500352500

Hedeker, D., & Gibbons, R. D. (1997). Application of random-effects pattern-mixture models for missing data in longitudinal studies. *Psychological Methods*, *2*(1), 64–78. https://doi.org/10.1037/1082-989X.2.1.64

Heinonen, E., & Nissen-Lie, H. A. (2020). The professional and personal characteristics of effective psychotherapists: A systematic review. *Psychotherapy Research*, *30*(4), 417–432. https://doi.org/10.1080/10503307.2019.1620366

Horvath, A. O., & Greenberg, L. S. (1989). Development and validation of the working alliance inventory. *Journal of Counseling Psychology*, *36*(2), 223. https://doi.org/10.1037/0022-0167.36.2.223

Johnson, E. L., Mutti, M.-F., Springham, N., & Xenophontes, I. (2016). Mentalizing after mentalization based treatment. *Mental Health and Social Inclusion*, *20*(1), 44–51. https://doi.org/10.1108/MHSI-11-2015-0042

Karterud, S. (2011). *Manual for mentaliseringsbasert psykoedukativ gruppeterapi (MBT-I)*. Gyldendal akademisk.

Karterud, S. (2012). *Manual for mentaliseringsbasert gruppeterapi (MBT-G)*. Gyldendal akademisk.

Karterud, S. (2015). *Mentalization-based group therapy (MBT-G): A theoretical, clinical, and research manual*. OUP.

Karterud, S., & Bateman, A. (2010). *Manual for mentaliseringsbasert terapi (MBT) og MBT vurderingsskala. Versjon individualterapi*. Gyldendal akademisk.

Karterud, S., Folmo, E., & Kongerslev, M. (2020). *Mentaliseringsbasert terapi MBT*. Gyldendal Akademisk.

Karterud, S., Pedersen, G., Engen, M., Johansen, M. S., Johansson, P. N., Schlüter, C., Urnes, O., Wilberg, T., & Bateman, A. W. (2013). The MBT adherence and competence scale (MBT-ACS): Development, structure and reliability. *Psychotherapy Research*, 23(6), 705–717. https://doi.org/10.1080/10503307.2012.708795

Kvarstein, E. H., Folmo, E., Antonsen, B. T., Normann-Eide, E., Pedersen, G., & Wilberg, T. (2020). Social cognition capacities as predictors of outcome in mentalization-based treatment (MBT). *Frontiers in Psychiatry*, 11. https://doi.org/10.3389/fpsyt.2020.00691

Kvarstein, E. H., Pedersen, G., Folmo, E., Urnes, Ø, Johansen, M. S., Hummelen, B., Wilberg, T., & Karterud, S. (2019). Mentalization-based treatment or psychodynamic treatment programmes for patients with borderline personality disorder–the impact of clinical severity. *Psychology and Psychotherapy: Theory, Research and Practice*, 92(1), 91–111. https://doi.org/10.1111/papt.12179

Kvarstein, E. H., Pedersen, G., Urnes, Ø, Hummelen, B., Wilberg, T., & Karterud, S. (2015). Changing from a traditional psychodynamic treatment programme to mentalization-based treatment for patients with borderline personality disorder–does it make a difference? *Psychology and Psychotherapy: Theory, Research and Practice*, 88(1), 71–86. https://doi.org/10.1111/papt.12036

Lemma, A., Target, M., & Fonagy, P. (2011). *Brief dynamic interpersonal therapy: A clinician's guide*. Oxford University Press.

Lonargáin, DÓ, Hodge, S., & Line, R. (2017). Service user experiences of mentalisation-based treatment for borderline personality disorder. *The Mental Health Review*, 22(1), 16–27. https://doi.org/10.1108/MHRJ-04-2016-0008

Möller, C., Karlgren, L., Sandell, A., Falkenström, F., & Philips, B. (2017). Mentalization-based therapy adherence and competence stimulates in-session mentalization in psychotherapy for borderline personality disorder with co-morbid substance dependence. *Psychotherapy Research*, 27(6), 749–765. https://doi.org/10.1080/10503307.2016.1158433

Morken, K. T. E., Binder, P.-E., Arefjord, N. M., & Karterud, S. W. (2019). Mentalization-based treatment from the patients' perspective–what ingredients do they emphasize? *Frontiers in Psychology*, 10. https://doi.org/10.3389/fpsyg.2019.01327

Mundt, J. C., Marks, I. M., Shear, M. K., & Greist, J. M. (2002). The work and social adjustment scale: A simple measure of impairment in functioning. *The British Journal of Psychiatry*, 180(5), 461–464. https://doi.org/10.1192/bjp.180.5.461

Muran, J. C., & Barber, J. P. (2011). In J. C. Muran & J. P. Barber (Eds.), *The therapeutic alliance: An evidence-based approach to practice and training* (pp. 44–62). New York, NY: Guilford Press.

Nissen-Lie, H. A., Havik, O. E., Høglend, P. A., Rønnestad, M. H., & Monsen, J. T. (2015). Patient and therapist perspectives on alliance development: Therapists' practice experiences as predictors. *Clinical Psychology & Psychotherapy*, 22(4), 317–327. https://doi.org/10.1002/cpp.1891

Nissen-Lie, H. A., Monsen, J. T., & Rønnestad, M. H. (2010). Therapist predictors of early patient-rated working alliance: A multilevel approach. *Psychotherapy Research*, 20(6), 627–646. https://doi.org/10.1080/10503307.2010.497633

Nissen-Lie, H. A., Rønnestad, M. H., Høglend, P. A., Havik, O. E., Solbakken, O. A., Stiles, T. C., & Monsen, J. T. (2017). Love yourself as a person, doubt yourself as a therapist? *Clinical Psychology & Psychotherapy*, 24(1), 48–60. https://doi.org/10.1002/cpp.1977

Normann-Eide, E., Antonsen, B. R. T., Kvarstein, E. H., Pedersen, G., Vaskinn, A., & Wilberg, T. (2019). Are impairments in theory of mind specific to borderline personality disorder? *Journal of Personality Disorders*, 1–15. https://doi.org/10.1521/pedi_2019_33_417

Ollila, P., Knekt, P., Heinonen, E., & Lindfors, O. (2016). Patients' pre-treatment interpersonal problems as predictors of therapeutic alliance in long-term psychodynamic psychotherapy. *Psychiatry Research*, 241, 110–117. https://doi.org/10.1016/j.psychres.2016.04.093

Pedersen, G., Arnevik, E. A., Hummelen, B., Walderhaug, E., & Wilberg, T. (2017). Psychometric properties of the severity indices of personality problems (SIPP) in two samples. *European Journal of Psychological Assessment*, 35. https://doi.org/10.1027/1015-5759/a000436

Pedersen, G., Hagtvet, K. A., & Karterud, S. (2007). Generalizability studies of the global assessment of functioning–split version. *Comprehensive Psychiatry*, 48(1), 88–94. https://doi.org/10.1016/j.comppsych.2006.03.008

Pedersen, G., & Karterud, S. (2004). Is SCL-90R helpful for the clinician in assessing DSM-IV symptom disorders? *Acta Psychiatrica Scandinavica*, 110(3), 215–224. https://doi.org/10.1111/j.1600-0447.2004.00321.x

Pedersen, G., Karterud, S., Hummelen, B., & Wilberg, T. (2013). The impact of extended longitudinal observation on the assessment of personality disorders. *Personality and Mental Health*, 7(4), 277–287. https://doi.org/10.1002/pmh.1234

Pedersen, G., Kvarstein, E., & Wilberg, T. (2017). The work and social adjustment scale: Psychometric properties and validity among males and females, and outpatients with and without personality disorders. *Personality and Mental Health*, 11(4), 215–228. https://doi.org/10.1002/pmh.1382

Pedersen, G., Urnes, Ø, Hummelen, B., Wilberg, T., & Kvarstein, E. (2018). Revised manual for the global assessment of functioning scale. *European Psychiatry*, 51, 16–19. https://doi.org/10.1016/j.eurpsy.2017.12.028

Perepletchikova, F., Treat, T. A., & Kazdin, A. E. (2007). Treatment integrity in psychotherapy research: Analysis of the studies and examination of the associated factors. *Journal of Consulting and Clinical Psychology*, 75(6), 829–841. https://doi.org/10.1037/0022-006X.75.6.829

Rossi, G., Debast, I., & Van Alphen, S. (2017). Measuring personality functioning in older adults: Construct validity of the severity indices of personality functioning–short form (SIPP-SF). *Aging & Mental Health*, 21(7), 703–711. https://doi.org/10.1080/13607863.2016.1154012

Saarni, S. I., Suvisaari, J., Sintonen, H., Koskinen, S., Härkänen, T., & Lönnqvist, J. (2007). The health-related quality-of-life impact of chronic conditions varied with age in general population. *Journal of Clinical Epidemiology*, 60(12), 1288. e1281–e1288. e1211. https://doi.org/10.1016/j.jclinepi.2007.03.004

Safran, J. D., & Muran, J. C. (2000). Resolving therapeutic alliance ruptures: Diversity and integration. *Journal of Clinical Psychology*, 56(2), 233–243. https://doi.org/10.1002/(SICI)1097-4679(200002)56:2<233::AID-JCLP9>3.0.CO;2-3

Sharp, C. (2014). The social–cognitive basis of BPD: A theory of hypermentalizing. In C. Sharp & J. L. Tackett (Eds.), *Handbook of borderline personality disorder in children and adolescents* (Vol. 1, pp. 211–225). New York: Springer. https://doi-org.ezproxy.uio.no/10.1007/978-1-4939-0591-1_15

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., & Dunbar, G. C. (1998). The mini-

international neuropsychiatric interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry*, *59*(Suppl 20 (22–33)).

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford university press.

Skodol, A. E., Shea, M. T., Yen, S., White, C. N., & Gunderson, J. G. (2010). Personality disorders and mood disorders: Perspectives on diagnosis and classification from studies of longitudinal course and familial associations. *Journal of Personality Disorders*, *24*(1), 83–108. https://doi.org/10.1521/pedi.2010.24.1.83

Soeteman, D. I., Verheul, R., & Busschbach, J. J. (2008). The burden of disease in personality disorders: Diagnosis-specific quality of life. *Journal of Personality Disorders*, *22*(3), 259–268. https://doi.org/10.1521/pedi.2008.22.3.259

Spinhoven, P., Giesen-Bloo, J., van Dyck, R., Kooiman, K., & Arntz, A. (2007). The therapeutic alliance in schema-focused therapy and transference-focused psychotherapy for borderline personality disorder. *Journal of Consulting and Clinical Psychology*, *75*(1), 104. https://doi.org/10.1037/0022-006X.75.1.104

Stiles, W. B., & Goldsmith, J. Z. (2010). The alliance over time. *The Therapeutic Alliance: An Evidence-Based Guide to Practice*, 44–62.

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. https://doi.org/10.1037/1040-3590.1.3.207

Tryon, G. S., Blackwell, S. C., & Hammel, E. F. (2008). The magnitude of client and therapist working alliance ratings. *Psychotherapy: Theory, Research, Practice, Training*, *45*(4), 546. https://doi.org/10.1037/a0014338

Vaskinn, A., Antonsen, B. T., Fretland, R. A., Dziobek, I., Sundet, K., & Wilberg, T. (2015). Theory of mind in women with borderline personality disorder or schizophrenia: Differences in overall ability and error patterns. *Frontiers in Psychology*, *6*, 1239. https://doi.org/10.3389/fpsyg.2015.01239

Volkert, J., Hauschild, S., & Taubner, S. (2019). Mentalization-based treatment for personality disorders: Efficacy, effectiveness, and new developments. *Current Psychiatry Reports*, *21*(4), 25. https://doi.org/10.1007/s11920-019-1012-5

Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work*. Routledge.

Young, J. E., Klosko, J. S., & Weishaar, M. E. (2006). *Schema therapy: A practitioner's guide*. Guilford Press.