FACULTY OF PSYCHOLOGY
AND EDUCATIONAL SCIENCES

Opening the black box of psychotherapy outcome: A mixed methods study on patients' perspectives

Melissa De Smet

Supervisor: Prof. Dr. Reitske Meganck

A dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of
Doctor of Psychology

Academic year 2018–2019

GHENT
UNIVERSITY

FACULTY OF PSYCHOLOGY
AND EDUCATIONAL SCIENCES

Opening the black box of psychotherapy outcome: A mixed methods study on patients' perspectives

GHENT
UNIVERSITY

Doctoral guidance committee

Supervisor

Prof. dr. Reitske Meganck (Ghent University, Faculty of psychology and educational sciences, Department of psychoanalysis and clinical consulting)

Guidance committee

Prof. dr. Mattias Desmet (Ghent University, Faculty of psychology and educational sciences, Department of psychoanalysis and clinical consulting)

Prof. dr. Geert Crombez (Ghent University, Faculty of psychology and educational sciences, Department of experimental clinical and health psychology)

Prof. dr. Ilse Derluyn (Ghent University, Faculty of psychology and educational sciences, Department of social work and social pedagogy)

During the fulfillment of this dissertation, Melissa De Smet was funded by the **Flemish Research Foundation (FWO)**

Examination board

Chair

Prof. dr. Peter Vlerick (Ghent University, Faculty of psychology and educational sciences, Department of personnel management, work and organizational management)

Exam committee

Prof. dr. Ignaas Devisch (Ghent University, Faculty of Medicine and Health Sciences, department of public health and primary care)

Prof. dr. Ellen Driessen (Radboud University, Radboud Behavioral Science Institute)

Prof. dr. Geert Van Hove (Ghent University, Faculty of psychology and educational sciences Department of special needs education)

Prof. dr. Tine Vervoort (Ghent University, Faculty of psychology and educational sciences, Department of experimental clinical and health psychology)

Nothing is objectively simple or complex except that man makes it so. This is the reason that philosophers and scientist who seek an objective denotation for 'simplicity' are bound to fail. There isn't any.

Simplicity is intimately related to understanding. Whenever a person cries out in relief: 'Aha! Now I understand what is going on', it indicates that what he previously experienced as a complex mess has suddenly made at least one quantum jump towards simplicity. The fact that often, upon reflection, the person may later find that he was wrong, that his solution does not hold, is (ir)relevant in this context.

Nehemiah Jordan, Themes in speculative psychology (p. 11, brackets added)[1]

[1] This citation grasps the central challenge that is inherently part of research on human complexity, in general and within this dissertation specifically. It explains that every attempt aimed at grasping complexity, for instance as expressed in patients' narratives, comes down to a necessary simplification in function of understanding. However, note that we have slightly adjusted the quote (see brackets) in order to stress the importance of continuously questioning the simplifications being made, or in other words, there is no definite end point in finding evidence or reaching understanding. Hence, after reaching the end of this dissertation, we ask the reader to return to this citation and continue the process.

Acknowledgements

This work comprises the written outcome of a years-long process in which I explored new grounds, often got stuck, and eventually found a language to express what I discovered. Before guiding the reader through this process and arriving at its outcome, I want to express a few words of gratitude to all those who have accompanied me along this journey.

First of all, I want to take the time to express my admiration for all the patients who have participated in this ambitious research project. Likewise, this would not have been possible without the help of the participating therapists. I want to thank each and every one of you for the openness to share your unique story and brave experiences.

The fulfilment of this work would not have been possible without the thoughtful guidance of my supervisor, Reitske Meganck. The value of a supervisor who has the capacity to think *with* (not for), *guide* (not control) and enable exploration (yet without you getting lost) is highly underestimated in the process of conducting a PhD. My ambitious and idealistic quest to "change the world of psychotherapy outcome research" would have not been translated into a feasible research project without you watching over it. Reitske, like most truly bright people, you do not get enough credit for all the things you accomplish.

I want to thank the members of the guidance committee, Mattias Desmet, Geert Crombez and Ilse Derluyn for the challenging (I must admit, sometimes frustrating) discussions. I am proud and thankful that we ultimately arrived at a shared understanding, transforming an initially complex idea into a coherent story.

A special thanks to all my colleagues, who I have become to see as *framily* along these intense years. Especially my RCT-sisters, who made the RCT a true life-altering experience. From sticking posters, to painting walls, to saliva samples and folding boxes, there is nothing we haven't shared by now. Thanks for putting up with me during stressful days and thanks for being your unique selves; from all of you I've learned a great deal. Femke, who has been a mentor, academic twin and bunny-loving friend from day one, your power to empower – *screw convention and modesty!* – has been of indispensable value to me as a person and researcher. I guess the famous *Femmel plant-book-coffee-store* will have to wait until we're finished changing the world of research, and I guess also *Albert Heijn* will not see our applications soon. I truly cannot wait for the projects that lay ahead and I certainly could not have pulled this one off without your ability to give voice and spark ideas. To Ufuoma, who as a real sister gets my randomness without thinking its random, thanks for the careful reading of my work, for taking care of me as a big sister and especially for letting me in as a friend. Rosa, I truly (and jealously) appreciate your authentic, open and honest way of being. Almost as much as your skill to make cocktails and getting the party started (let's never change that tradition!). Vicky, forever my kayak-sister, your ability to calm me down in a spinning boat prove you not only study but also possess the magical effect of *transference*. Let's kayak again soon, I promise to stick to your instructions this time. Goedele, thanks for the numerous relieving talks that put life and "saliva-induced" frustration into perspective. Your humour and sharp mind offered a welcome break from reality. Kimberly, *Kaap* who's the bravest and most dangerous of us all, I truly admire your turbo way of being in life which dazzles me every time. No doubt, you will achieve what you aim for.

Thank you, Joachim, for being part of the magnificent team-weekend and an always available office-neighbour. I will never forget your help with grasping the Graph in a pedagogically appropriate manner, nor the dancing and singing involving trips to Jupiter, a Lacanian frog and mozzarella plant. To Jasper, thanks for sharing an office and tolerating my plants, love for rice crackers and other characteristics. For your rosé-wisdom, especially, but maybe most of all for the rosé itself: *alles komt goed*, as we know. Thanks to Junior for the hard work (code name for office chair-racing down the hallway) during the Holidays and weekends. Els, your smile is a blessing on days in which the printer breaks down (again), the coffee runs out and administrational chaos kicks in. Thanks for the work you always do like only you can do. Bless you David for bringing *het écht Antwaarps theater* to the department, *amai maaine frak gast*. Thanks to Liza for sharing unforgettable nights at the Beethoven-bed-bug hotel, and Juri for sharing a special admiration for the Belgian exclusive architecture. You will make theory great again (with your inspiring drawings)! Thanks to Lisanne, Emma, Nienke and Bart for sparking the department with fresh enthusiasm, motivation and bad jokes (yes you Bart). Thanks to Amar for the laughing-sessions in the hallway and welcome motivational quotes (*run-mellie-run*). Thanks to all other colleagues, for showing me the way, and/or sharing too many drinks at Vakgroep-weekends: Dries, Eline, Ellen, Evi, Abe, Gregory, Virginie, Kaatje, Wendy, Wim, Els, Eli, and Ruth (forever the RCT big sister). A special thanks to Shana, for guiding me through the last months of this work, the careful editing and your motivational cheers.

Thanks to the department for being a warm nest I could return to during my journey that took me to many places, wandering but never lost. Thanks especially to Stijn for putting trust in our work on the therapy project. To Paul for his inspiring lectures that were the motivation to start this PhD adventure in the first place. To Mattias for not taking convention for granted and daring to ask the complexity questions. To Filip, for your kindness and empowering belief in us young ones, you are truly missed. To Hubert, for your book.

The writing of this work would certainly not have been accomplished without my dear fellow-members of the one-and-only Pinocchio writing group. Kaat, the green power ranger, your skill to search and find the extraordinary at *tweedehands.be* is truly admiring, thanks for the late-night writing sessions, *knallen en pompen,* you're almost there and yes you will make it! Céline, time to pick up on those marathon trainings again, no excuses now! Thanks to Rosa and Patricio for adding a scent of good-life (in the form of food, drinks and music) to the work atmosphere, and Yannice and Sofie for showing how concentrated writing works.

To all the people I've met during this journey who have inspired my thinking, research and writing. Camilla von Below and Andrzej Werbart, especially, as well as the rest of house 8 and 14 at Stockholm University: curling and bird watching have certainly been life-altering experiences, and I still miss *fika* on Thursdays. Special thanks to John McLeod and the department of psychotherapy at Oslo University, where it all began, and the critical minds at the SPR community who introduced me to this field. Special thanks to Bill Stiles, Giulio de Felice, Omar Gelo, Felicitas Rost, Stig Poulsen, Imke Ahlers and so many others for co-hosting panels and discussions and striving for better psychotherapy research and patient health care.

From the heart, I want to thank some of the most important people in my life:

My friends, who probably cherish the idle hope for me to obtain a “normal” work-life balance now. I hope to convince you of the important idiosyncrasy in that experience and thus to assure you there’s no need to worry. Shani and Bart, *Shaba*, who by now should know that I will follow them around the globe wherever they may go. Thank you for always offering me a house to come home to. Stefanie and Dieter, thanks for the best Q-time and being the best roomie, and to all other members of the Table of Honour, Maxine, Evelien and Arum for the wonderful food-and-drinks traditions.

To Anna, for understanding without having to say a word and for the soothing effect that only your words have.

To Julie, soon to be doctor Rosseel, for seeing beauty in the ugly, for being so strong in an unfair world, for being you.

To Joni for being perhaps the only one who can both handle and meet my exceptional dance moves.

To Tine, Christel, Jeroen and Nathalie for still being there since back in the days, let’s meet more often, starting soon.

To my parents, for their everlasting support and trust in my ambitious plans, to follow me no matter where my dreams or heart may take me and to always belief that things we’ll be alright. I could not have pulled this off without your encouragement, patience, assistance and genuine enthusiasm.

To my brother, Michaël, who will always (literally) outrun me with his ambitious plans and who I’m most proud of. Only you do what you do, and all at once with the same scrutiny. To Katherine, for being his calm other half.

To my grandfather, who’s critical ideas and passionate beliefs have certainly affected me.

To Derek, *Dre*, a special award for the support and distraction during the intense final months of this PhD life. I am very happy you happened to *happn* and I look forward to the many adventures of *double duts*.

To Italy, for the food, wine, culture and beautiful landscapes; for showing me the value of *dolce far niente*.

Abbraccio.

Melissa, 28 June 2019

Content

I. GENERAL INTRODUCTION 9

THE GAP BETWEEN RESEARCH AND PRACTICE, AND THE PIVOTAL ROLE OF PSYCHOTHERAPY OUTCOME RESEARCH 10

PSYCHOTHERAPY OUTCOME: DIFFICULT TO DEFINE, EASIER TO MEASURE? 13

CLINICAL SIGNIFICANCE: A STATISTICAL ANSWER TO CLINICAL MEANINGFULNESS 14

THE "BLACK BOX" OF PSYCHOTHERAPY OUTCOME 16

"Clinically significant" but arbitrary? *16*

"Clinically significant change", but how? *17*

OBJECTIVES 18

THE PATIENTS' PERSPECTIVES: DEEPENING THE UNDERSTANDING OF CHANGE AND THERAPY 19

OPENING THE "BLACK BOX": A MULTI-LAYERED ANALYSIS OF CHANGE WITHIN PSYCHOTHERAPY OUTCOME RESEARCH 21

METHOD 22

EXPLANATORY SEQUENTIAL MIXED METHODS STUDY 23

First-person perspective: giving voice to the patients' narrative experiences *27*

Qualitative, data-driven analyses departing from a contextual perspective *28*

Situating the present dissertation within the process-outcome field: A focus on "change" *29*

OVERVIEW OF THE CONDUCTED STUDIES AND CHAPTERS 29

Pilot study: Clinical practice calling for research evidence. *30*

RCT: Standard outcome research in search of clinical meaningfulness. *31*

REFERENCES 35

II. UNDERSTANDING LONG-TERM OUTCOME FROM THE PATIENTS' PERSPECTIVE: A MIXED METHOD NATURALISTIC STUDY ON INPATIENT PSYCHOTHERAPY 43

INTRODUCTION 44

METHOD 47

RESULTS 55

DISCUSSION 65

CONCLUSION 70

REFERENCES 70

III. WHAT "GOOD OUTCOME" MEANS TO PATIENTS: UNDERSTANDING RECOVERY AND IMPROVEMENT IN PSYCHOTHERAPY FOR MAJOR DEPRESSION FROM A MIXED-METHODS PERSPECTIVE 76

INTRODUCTION 77

METHOD 81

RESULTS 91

DISCUSSION 101

CONCLUSION 110

REFERENCES 111

IV. CHANGE PROCESSES UNDERLYING "GOOD OUTCOME": A MIXED METHODS STUDY ON RECOVERED AND IMPROVED PATIENTS' EXPERIENCES IN PSYCHOTHERAPY FOR MAJOR DEPRESSION 118

INTRODUCTION 119

METHOD 122

RESULTS 130

DISCUSSION 138

CONCLUSION 146

References ... 146

V. NO CHANGE? A GROUNDED THEORY ANALYSIS OF DEPRESSED PATIENTS' PERSPECTIVES ON NON-IMPROVEMENT IN PSYCHOTHERAPY ... 152

Introduction ... 153
Method ... 157
Results ... 166
Discussion ... 177
Conclusion ... 189
References ... 189

VI. MAKING SENSE OF DETERIORATION IN PSYCHOTHERAPY: AN INTERPRETATIVE PHENOMENOLOGICAL ANALYSIS OF PATIENTS' PERSPECTIVES ... 198

Introduction ... 199
Method ... 201
Results ... 207
Discussion ... 218
Conclusion ... 224
References ... 224

VII. GENERAL DISCUSSION ... 229

Main study findings ... 230
Integrative conceptual model of psychotherapy outcome: Synthesis of patients' perspectives ... 235
"Outcome" as an ongoing, non-linear process characterized by individual differences ... *236*
"Outcome" as a holistic experience of multi-dimensional change ... *237*
Multiple change processes – in and outside therapy – interact over the course of time. ... *242*
Psychotherapy outcome does not necessarily reflect experienced psychotherapy effects ... *248*
The clinical (un)meaningfulness of statistical outcome classification ... 248
"Clinical significance" versus the multiplicity and nuance in patients' experiences ... *250*
"Clinical significance" versus cross-sample similarities and within-sample differences ... *251*
Ergo, there's no going back after opening the "black box" ... 254
Towards a "first-person" understanding of clinically significant change ... 259
Implications and challenges for outcome research, clinical practice and policy ... 261
Outcome research departing from a broader interpretation of "clinical significance" ... *261*
What counts as "evidence" and the implications for health care decisions ... *267*
Research developments becoming clinical applications ... *268*
Policy decisions, the tension between economical and clinical needs ... *270*
Strengths, limitations and future directions ... 272

CONCLUSION ... 281

References ... 282

ENGLISH SUMMARY ... 293

NEDERLANDSTALIGE SAMENVATTING ... 298

APPENDICES ... 304

Appendix 1. Study 1. Clinical description, diagnoses and total scores (OQ-45) per patient. ... 304
Appendix 2. Study 4. Case-specific findings: summary ... 308
Appendix 3. Data Storage Fact Sheet ... 311

I. General introduction

Psychotherapy research is concerned with providing research findings that can be useful in the organization of mental health care. However, the translation of research findings to clinical practice is complex; specifically, the clinical generalizability of general outcome research is debated frequently in scientific literature and by clinical professionals. This dissertation is focused at the clinical relevance *within* research findings, with the aim of elucidating clinical meaningfulness in general findings that can be useful for this translation. For this purpose, the aim is to open up general statistical categories that are dominant in outcome research, by focusing on the patient perspectives on which these categories are build up. By scrutinizing both the clinical and methodological relevance of general categories, this dissertation aims to open the "black box" of outcome in psychotherapy research.

This general introduction starts with a brief outline of the historical tension between outcome research and clinical practice. Subsequently, we turn to the conceptualization and measurement of psychotherapy outcome, and the attempts that have been made to provide clinically relevant outcome findings. Finally, we introduce the central research objectives this dissertation aims to address, describe the applied methodology and present the five chapters that comprise the four mixed methods studies conducted in this project.

The gap between research and practice, and the pivotal role of psychotherapy outcome research

Dissemination of research findings into clinical practice is of central importance to researchers, policy members, clinicians and patients. This is known as Evidence-Based Practice (EBP), which implies "the integration of the best available research with clinical expertise in the context of patient characteristics, culture, and preferences." (APA, 2006, p. 271). "The best available research" is generally regarded as sample-based outcome research, in which the efficacy of treatments is systematically compared (Wampold & Imel, 2015). This design yields a number of stringent methodological requirements (see Consort, 2010[2]). In particular, standardization of the research setting (e.g., eligibility criteria for the inclusion of participants, use of treatment protocols), random allocation of patients either to a treatment or control group, and pre-registration of the primary and secondary outcome measures being used to evaluate outcome. The efficacy of the treatment is defined by means of statistical analysis comparing the treatment and control group with regard to their mean-level of change. This "gold standard" rigorous methodology, known as randomized controlled or causal effects research (hereafter referred to as RCTs), is located at the top-side of the hierarchy of evidence (Jadad & Enkin, 2007; Lambert, 2013). However, while this stringency is seen as an advantage in terms of methodological rigor, it brings about a major problem for the possibility of translating findings to clinical practice – which context and characteristics are much less clear-cut (Westen, Novotny, & Thompson-Brenner, 2004).

Formally, this debate entails the tension between internal validity on the one hand and external and ecological validity on the other. While the first allows for causal claims regarding the general efficacy of a certain treatment, the others are concerned with whether these findings can also be translated into contexts outside of the research setting and the applicability to actual

[2] http://www.consort-statement.org

clinical care (see Goldfried & Wolfe, 1996 and Westen et al., 2004 for a critical discussion). Beyond the problem of external validity/generalizability, however, critics also object to the findings themselves; that they are not clinically sound or internally valid, as they are based on *mean* difference between groups or categories of people and hence cannot provide insight into clinically representative processes or mechanisms (Dattilio, Edwards, & Fishman, 2010).

To streamline the translation of research evidence into clinical applications, currently the focus is on training of clinical professionals, distribution of and adherence to protocols, and increasing the accessibility of empirical findings (see Castonguay, Barkham, Lutz & McAleavey, 2013, Jadad & Enkin, 2007, and McLeod, 2016 for an overview and detailed discussion). However, while these incentives play a role in *dissemination* of EBT, they do not yet aid the translation from stringent methodology to complex clinical practice. Therefore, a pivotal role is reserved for psychotherapy researchers to provide the field with clinically *relevant* and useful findings that enable continuous improvement of mental health care for patients. Particularly, as the empirical standards of outcome research have been largely adopted by governments and policymakers as a warranty for qualitative health care, a direct link exists between outcome research and policy decisions that impact the organization of clinical care (cf. McLeod, 2016).

In the present dissertation we therefore address the clinical meaningfulness of psychotherapy outcome findings. Currently, "clinical meaningfulness" is considered in terms of "clinical significance" of measured outcome, which is the statistical indication of the clinical relevance of measured change (Jacobson & Truax, 1991; Lambert, Hansen & Bauer, 2008). The development of this index sprung from the same observation that we make in this dissertation, namely the need to incorporate relevance for clinical practice into general (mean-level) outcome research. This way, the clinical significance turned an important focus to *patient-centeredness* of statistical outcome (Jacobson, Folette & Revenstorf, 1984) and

therefore serves as the particular starting point of the present dissertation. However, whereas "clinical significance" focuses on the relevance of general (sample-based) outcome by providing an indication of the variation within the samples being compared, it does not necessarily capture the question of *usefulness* of findings with regard to actual clinical applications (cf. the ecological validity; Goldfried & Wolfe, 1996). Hence, also the extent to which outcome findings enable policy members and clinicians to draw conclusions that, in a second step, can inform (the organization of) clinical practice with individual patients will be discussed (cf. McLeod, 2016).

The issues that we address and the avenues we propose are in the first place focused on the clinical relevance of RCT-based outcome, given their status as best available method. However, the issues are not limited to RCT-contexts, as they concern the broader range of outcome research relying on statistical outcome measures. Therefore, throughout this dissertation, we will often refer to "standard outcome research" to refer to studies that rely on statistical procedures to define the outcome of psychotherapy. In this, we explicitly focus on pre-post evaluation of outcome using standardized (pre-defined) quantitative outcome measures, that are concretely used to evaluate patient change in therapy research (Greenberg & Pinsof, 1986). As the goal is to understand usefulness towards clinical practice, we do not focus on psychometric nor statistical features, nor elaborate on technical problems of RCTs, as these are discussed at length elsewhere (e.g., Wampold & Imel, 2015). Rather, we elaborate on the patient perspectives entailed in the findings, which will indeed be used as a basis for clinical practice. This way, we present a representative case of the research findings that are being relied on to inform policy decisions.[3]

[3] It is interesting to note that similar concerns about the statistical evaluation of treatment efficacy have been voiced in other health care domains, for instance the field of health psychology and particularly the treatment of chronic pain. Although these concerns about how clinical effectiveness can best be regarded and studied have been answered using a somewhat different approach, they reflect similar issues in the translation of standardized research and gold standard methodology to the fields that deal with less clear-cut phenomena, like mental health, psychotherapy and chronic pain complaints (see for instance Morley, Williams & Eccleston, 2013).

Psychotherapy outcome: Difficult to define, easier to measure?

Despite the strong emphasis on outcome measurement in psychology, a clear definition of what "psychotherapy outcome" *is*, is lacking (Evans & Fletcher, 2013; Rodgers & Elliott, 2013). Defining the outcome of an intervention inevitably depends on a value judgement, based on an idea of mental wellbeing and an underlying theory of what psychotherapy should achieve (Evans & Fletcher, 2013; Strupp & Hadley, 1977). Indeed, the definition tends to vary for different theoretical orientations and from perspectives (cf. Levitt, Stanley, Frankel & Raina, 2005). For instance, patients, clinicians, society and peers all hold somewhat different ideas of what satisfying outcome and therapy looks like (e.g., society being concerned with people being able to go work may differ from patients need for self-exploration; Strupp & Hadley, 1977).

Despite the lack of clear understanding or consensus of what outcome *is*, psychotherapy researchers – concerned with establishing a comparative ground of evidence and coherent research field – have relied strongly on a *methodological definition* of outcome in adherence to gold standard research stringency (see Braakmann, 2015; Wampold & Imel, 2015 for an overview of the field's history). Concretely, this implies that *statistical* outcome[4] is currently prioritized over other – more value-based and orientation-specific – conceptualizations (Braakmann, 2015; Greenberg & Pinsof, 1986; Hill & Lambert, 2004; Lambert, 2013; Levitt et al., 2005).

Generally, three categories of statistical tests can be applied to evaluate psychotherapy outcome. In fact, it is generally recommended that all three are used to evaluate the consequences of interventions in efficacy and effectiveness studies (Lambert et al., 2008; Westen et al., 2004). "Statistical significance" indicates that a mean difference is larger than what could have been expected by mere chance, which supports the interpretation that the difference is caused by a treatment intervention rather than by the mere passage of time or

[4] This generally refers to the pre-post measurement of patient change, predominantly symptom-based (see Greenberg & Pinsof, 1986; Hill & Lambert, 2004; Lambert, 2013)

circumstances. “Effect sizes” provide an indication of the magnitude of those mean differences (i.e., the strength of the observed effect), which is held as indicator of how well the treatments work. Both statistical significance and effect sizes regard *between-group* variation (i.e., between a treatment-group and a control-group). As a third statistical indicator, *clinical significance* was however proposed to shed light on the *within*-sample variation, as general (mean-level) statistical significance may be clinically *meaningless* with respect to the measured changes of the individual patients within the sample (Lambert et al., 2008; Lambert, 2013). In the present dissertation we particularly focus on the method of clinical significance as developed by Jacobson and colleagues (1984; 1991), who argued that neither statistical significance nor effect sizes actually grasped treatment *efficacy*, as “questions regarding the efficacy of psychotherapy refer to the benefits derived from it, its potency, its impact on clients, or its ability to make a difference in peoples’ lives” (Jacobson & Truax, 1991, p.12). In what follows, we will briefly outline the answer they put forward to do so.

Clinical significance: A statistical answer to clinical meaningfulness

Different approaches have been developed to indicate the clinical relevance or practical implication of outcome findings (e.g., “social validity” which relied on subjective evaluation by significant others, Kazdin, 1977; see Ogles, Lunnen & Bonesteel, 2001; Ogles, 2013 for a brief history), yet Jacobson, Folette and Revenstorf were the first to suggest a systematic statistical method aiming at consensus in the field (Lambert et al., 2008). They argued that:

> ‘the field needs *agreed-upon conventions* for determining what constitutes improvement in an individual, and, if possible, *general consensus* as to what is meant by clinical significance. These conventions should be applicable to a *wide variety of clinical problems*, and they should be *objective, relatively free of bias,* and sound from both a psychometric and a clinical perspective. Without such conventions these data will remain an afterthought, a post hoc exploration acting as a side-show to the main event which continues to be the group significance test. Conventions protect the field from capitalization on chance, and they guard against the tendency for an investigator to use self-serving criteria. Moreover, currently utilized criteria vary so widely from one study to another that there is little basis for comparison. It seems clear that if reports highlighting variability

> and clinical significance are to emerge as primary data for evaluating the effects of psychotherapy, the field needs an agreed-upon basis for deciding whether a client warrants being classified as "improved" and when that improvement is clinically significant.' (Jacobson et al., 1984; p.338-339, emphasis added).

More in particular, Jacobson et al. (1984) presented a two-fold evaluation of "clinical significance": statistically reliable change pre-to-post therapy (i.e., pass the reliable change index, RCI) and evolution from dysfunctional to functional distribution (i.e., pass the predefined cut-off)[5]. Using this twofold evaluation, a patient can be marked as "Recovered" (i.e., reliable change and patient evolved from dysfunctional to functional range by passing the cut-off)[6], "Improved" (i.e., reliable change but the patient remains within the range of the clinical population), "Unchanged" (the RCI criterion was not passed), or "Deteriorated" (the RCI criterion was passed in the negative direction, patient moves further down the dysfunctional range). With this definition, Jacobson et al. (1984) envisioned patients' outcome on a continuum, and dividing patients' functioning (dichotomously) into a functional and dysfunctional population, conventionalized the interpretation of outcome in reference to normative data and predefined cut-offs for a particular measure being used (Jacobson et al., 1984). Since then, this method – which was later refined by Jacobson and Truax (the "JT method"; 1991) – became the definition of choice to indicate the clinical meaningfulness of outcome findings and the most-often used method in psychotherapy research to date (Atkins, Bedics, McGlinchey, & Beauchaine, 2005; Lambert & Ogles, 2009).[7]

[5] In concrete, the RCI is calculated by dividing the difference between the observed post-test and pre-test scores by the standard error of differences (Jacobson & Truax, 1991). The size of the standard error of measurement depends on the reliability of the outcome measure. Hence, the more reliable the instrument, the smaller the resulting standard error and thus the smaller the observed change between pre- and post-test score required to achieve a statistically reliable change. As an RCI larger than +/- 1.96 would be unlikely to occur ($p < .05$) without actual change, it identifies reliable improvement or deterioration. Secondly, a cut-off point between a patient (dysfunctional) and a nonpatient (functional) population is estimated. Three different cut-off scores are calculated depending on the instruments' available normative data for a patient and/or nonpatient population. (see Jacobson & Truax, 1991; see Lambert et al., 2008 for a concrete example and instructions).

[6] Recovery and clinical significance are often used interchangeably (Lambert et al., 2008), in the present dissertation we predominantly utilize the term "recovery", unless indicated otherwise (e.g., in study 1).

[7] Other means have been developed as an answer to shortcomings of the JT method; a technical discussion of the various existing methodologies for calculating clinical significance is however beyond the scope of this dissertation (see Atkins et al., 2005;

The "black box" of psychotherapy outcome

The development of clinical significance meant an important contribution to the statistical evaluation of psychotherapy outcome, as it meant a shift towards an individual-patient indication of the relevance of the observed pre-to-post changes which offered important information on the variation underlying mean-level treatment effects (see Jacobson et al., 1984). Nonetheless, several questions remain regarding the understanding of psychotherapy outcome and the interpretation of clinically significant outcome findings that call for further conceptual and methodological developments (Kazdin, 2008; McLeod, 2011). In this dissertation, we focus on two specific questions that remain unclear in the "black box" of psychotherapy outcome. Based on these questions, we derive three general consecutive research objectives for the current project, which will be outlined below.

"Clinically significant" but arbitrary?

A first remaining question that puzzles the understanding of outcome and its clinical meaning, is what outcome measures and "clinical significance" signal in light of patients' experiences. Specifically, given the strong dependence on the concrete instrument that is used to evaluate pre-to-post change (cf., Jacobson et al., 1984), the JT definition is susceptible to the difficulties that are voiced for standard outcome measurement more generally (Lambert et al., 2008). For instance, one-dimensional, predominantly symptom-based measures, are most-often used to evaluate outcome (Braakman, 2015; Levitt et al., 2005), although research shows that symptoms alone are an insufficient indication of patient change (McLeod, 2011; 2013; Zimmerman et al., 2012) and the complex problems patients are confronted with (Hill, Chui & Baumann, 2013). While in recent years, researchers increasingly use multiple measures (encompassing different levels of functioning, e.g., interpersonal relations) and respondents or

Maassen, 2000; McGlinchey, Atkins & Jacobson, 2002; Ogles et al., 2001). Here we rely on the JT method of clinical significance as it the most widely used measure, but the issues addressed are probably not limited to this specific case (see Lambert et al., 2008; Wise, 2004; Zimmerman et al., 2012).

methods (e.g., therapist and family members, observational and self-report measures) to provide a more diverse picture of outcome (Hill & Lambert, 2004), interpreting discordance among different measures remains an unsolved problem (see Atkins et al., 2005; De Los Reyes et al., 2011; Ogles et al., 2001; Ogles, 2013). Most importantly, outcome measures are considered "arbitrary metrics" as the scores on a certain measure do not represent the underlying (real-life) construct, as "a *measure of* something, is not the something itself." (Evans & Fletcher, 2013, p.32). In other words: scores on an outcome measure do not necessarily identify the extent to which patients have been helped in their *daily lives,* which requires real-life contextualization (Slife, 2004).

Consequently, it is possible that a patient scores within the functional range after therapy and will thus be considered improved in a clinically significant way, while it does not reflect the person's functioning in everyday life (Kazdin, 2006). This is highly problematic with regard to the clinical meaningfulness and *usefulness* of findings, as a statistical indicator of change may obscure actual patient needs that are of inherent value to clinical practice. Thus, to be of clinical meaning, it is argued that "there needs to be further conceptual clarity about what is precisely meant by clinical significance vis-à-vis the impact of therapy on clients' lives" (Atkins et al., 2005, p. 987) and consequently, more insight into how outcome and the effects of psychotherapy are experienced and perceived by patients themselves (see Carey et al., 2006).

"Clinically significant change", but how?

A second question that remains unanswered when relying on standard pre-post outcome evaluation is *how and why* the observed changes came about (Evans & Fletcher, 2013; Kazdin, 2009). While a robust corpus of evidence supports the general efficacy of psychotherapy over "doing nothing" (Munder et al., 2019), attention has turned to the understanding of the underlying trajectory or operating mechanisms (Elliott, 2010). In the field of psychotherapy research, the study of outcome and processes have mainly been an *either-or* endeavour

(Greenberg & Pinsof, 1986). This strict distinction between changes occurring outside of therapy (cf. evaluated pre-post) and *in*-session processes have been criticized as an "arbitrary dichotomy", resulting in a narrowed understanding of therapy effects which does not accord to clinical reality (see Kiesler, 1966; 1971; 1973). In response to the limits of studying processes regardless of outcome, and the evaluation of outcome without understanding the processes, the past decades have witnessed a shift toward emphasizing the study of *change,* which reduced the process-outcome dichotomy (Greenberg & Pinsof, 1986).

Hitherto, an evidence-based understanding of the mechanisms underlying psychotherapy outcome is lacking, however (Cuijpers, Reijnders & Huibers, 2019; Kazdin, 2009). Indeed, the linear cause-effect relationships underlying pre-post outcome research are deemed too simplistic to grasp how psychotherapy operates, and thus, particularly causal research designs like RCTs are being challenged to provide a plausible explanation for the nature of the observed relationship. Proving general efficacy *apart from* a minimal understanding of the underlying or occurring in-treatment change processes has become untenable (Elliott, 2010). Therefore, in order to inform clinical practice as well as to build further knowledge regarding how and why psychotherapy works (rather than proving *that* it works), psychotherapy research into treatment outcome should engage in examining the change processes or trajectory underlying pre-post outcome findings, beyond its "clinical significance" (Elliott, 2010; Evans & Fletcher, 2013).

Objectives

In the present dissertation we depart from the above-outlined observation that despite important developments to provide clinically meaningful outcome findings for individual-patient change, pivotal questions remain that consequently call for further conceptual and methodological developments to provide policy and clinical care with useful findings. Particularly, what "clinically significant change" implies in light of patients real-life and therapeutic experiences

remains unclear, with preliminary findings pointing at different directions, hence masking the *actual* clinical meaning (Atkins et al., 2005; Kazdin, 2001a). Second, general efficacy does not provide insight into the nature of the measured changes, hence leaving clinical care with the finding that therapy works (in general) yet without a clue how (Elliott, 2010; Kazdin, 2009). As these questions are most loudly voiced for the influential standardized controlled efficacy studies, we will particularly address them *within* the context of an actual RCT.

Specifically, building onto the earlier developments to provide clinically relevant outcome findings in terms of individual-patient change, we continue this endeavour by extending the understanding of both "outcome" and its underlying change processes by inquiring patients' perspectives. Specifically, we (1) scrutinize the changes that patients have experienced over the course of psychotherapy, in order to grasp how outcome can be understood from patients' perspectives; and (2) elucidate the processes to which patients ascribe meaningful change themselves, which allows for a broader understanding of how and why changes come about over the course of psychotherapy. Subsequently, relying on patients experiences as they are presented *within* the pre-defined statistical categories we (3) examine how well the predominantly used statistical evaluation of psychotherapy outcome, in terms of clinical significance as proposed by Jacobson and Truax (1991), can capture clinically meaningful change from patients' perspectives on their own functioning after therapy. Ultimately, this allows us to reflect on the usefulness of standard outcome research for the transference of findings to clinical care, which we will return to in the general discussion.

The patients' perspectives: Deepening the understanding of change *and* therapy

The observations that we presented have been raised before, and fueled a call for qualitative research in psychology, to entertain a more explorative method to examine the perspectives of patients (Braakmann, 2015). While for a long time, patients' voices have been neglected in the study of processes and outcome, slowly, the awareness grew that especially

these insider-perspectives entail a wealth of clinically relevant knowledge (Carey et al., 2007). As patients are "the site of change" (Greenberg, 1991, p.10) their perspective on treatment outcome and the processes that lead to change provide clinicians with important insights on the helpful or hindering role of treatment (McLeod, 2013). Qualitative inquiry especially allows for in-depth examination and contextualization within lived experiences (Ponterotto, 2005) and is therefore strongly appealing to clinicians (see McLeod, 2013; 2016). For the same reason, it is particularly considered of added clinical value to increase the yield of randomized controlled studies (see Datillio et al., 2010; Midgley, Ansaldo & Target, 2014). Moreover, as qualitative research especially allows one to explore phenomena independently of predefined definitions, it is deemed especially suitable to shed light on underexplored or ill-understood phenomena like change processes and outcome of psychotherapy (Rodgers & Elliott, 2013). As noted by Rodgers (2003) this approach particularly allows grasping *the reasons behind the observed findings*. In that line, the present dissertation puts the patient perspective central, to shed light on the complex matter of psychotherapy outcome and change processes by means of in-depth qualitative analysis.

More specifically, the first objective of this dissertation advances the (steadily) growing line of qualitative research, aiming to increase the understanding of psychotherapy outcome departing from a first-person patient perspective (see Levitt et al., 2016; McLeod, 2013). The second objective builds onto the strand of change process research concerned with "identifying, describing, explaining, and predicting the effects of the processes that bring about therapeutic change'' (Greenberg, 1999). More particularly, the "qualitative helpful factors" approach which relies on patients' experiences to shed light on helping and hindering elements of psychotherapy (see Elliott, 2010; and Timulak, 2007; 2010). Hence, apart from addressing the methodological question regarding the *study* of psychotherapy outcome, this research project also serves a *clinical* and *conceptual* cause, aiming at deeper understanding of the complex phenomena of

outcome and change processes.

Opening the "black box": A multi-layered analysis of change *within* psychotherapy outcome research

Addressing the above-outlined objectives *within* the context of an actual RCT allows a systematical examination of the third objective, which is the evaluation of the clinical usefulness of the statistical indices of "clinical significance" (by Jacobson & Truax, 1991) as a means to provide a clinically relevant indication of patient change. Particularly, we examine the experiences of those *patient-participants* (i.e., the randomized participants; cf. Truijens, 2016) as they are present *in* the RCT and *within* the general statistical categories of clinical significance. Specifically, we depart from the classifications of "recovery" and "improvement", which are typically regarded as "good" outcome, and "no change" and "deterioration", which are typically identified as "poor" outcome (cf. Jacobson & Truax, 1991; McElvaney & Timulak, 2013).

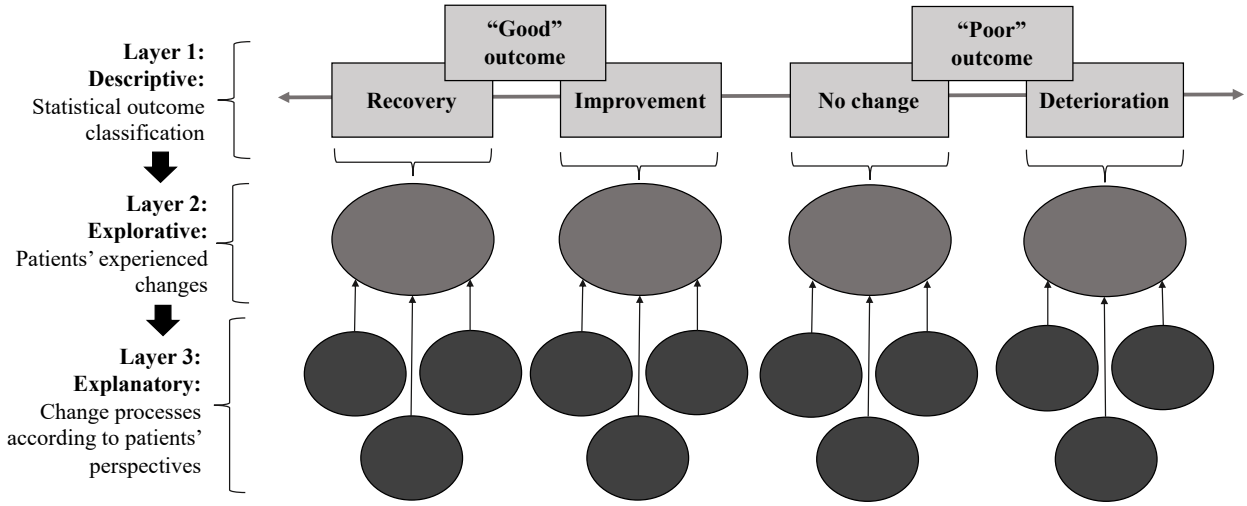
Conceptually, this can be understood as a *multi-layered* analysis to the study of psychotherapy outcome (broader: change). Departing from the *descriptive* categories identified by means of clinical significance indices (Jacobson & Truax, 1991), an *exploration* of patients' experienced changes (i.e., aiming at understanding their "outcome" experiences) and the factors they believe have influenced these changes (i.e., aiming at tentative *explanations* and theory-building) are studied subsequently (see also Kazdin, 2001b). This procedure is depicted in Figure A and will be described in more detail in the method section.

Importantly, when expanding the standard statistical evaluation of outcome by means of in-depth exploration of patients' experiences, we do not aim at a mere comparison of both outcome evaluations. Given the distinctive nature of both approaches, it would be stating the obvious that the latter will provide "richer" or different insights comparing to what is observed by means statistical measurement alone. Rather, the aim is to explore *what* the patients'

perspectives contribute to the understanding of change (rather than proving *that* it will contribute something), in order to deepen the understanding of outcome as well as the underlying change processes. Importantly, by combining both forms of evaluation (or perspectives) in *one and the same research*, in a second step, it allows us to reflect on whether the statistical indices of clinical significance, which aim at providing a *patient-centred* indication of the relevance of pre-post changes (Jacobson & Truax, 1991) live up to that task (Kazdin, 1999; 2001a).

Thus, the present dissertation addresses both a clinical and methodological concern when attempting to open the "black box" of psychotherapy outcome and increasing the clinical meaningfulness of outcome findings by extending standard statistical outcome evaluation with patients' subjective experiences. The applied methods and procedures deployed to do so are outlined hereafter.

Figure A. Multi-layered analysis of change in psychotherapy as applied in this research



Method

As our central aim concerns the exploration of patients' particular experiences of their therapy outcome and therapeutic process leading up to this, in relation to the statistical evaluation of

clinical significance, a mixed method (quan → QUAL) approach was considered most suited. Mixed methodology represents the "third research paradigm" and is particularly suitable to create a dialogue between distinct perspectives and integrates both quantitative and qualitative methods into one research (Tashakkori & Teddlie, 2003). Midgley, Target and Smith (2006) speak in this regard of a "Rashomon experience" (referring to the Japanese tale and famous film of Akira Kurosawa) or "triangulation" that occurs when an extra perspective (e.g., the patient) is added to the existing conceptualization (e.g., the predefined statistical definition), as it has the ability to induce doubt, or create new meaning. Different approaches to mixed methods research, however, exist. Given the central aim to understand patients' experiences within their idiosyncratic context, in the present dissertation the main emphasis is on the qualitative strand of the mixed design. This adheres to an interpretative approach to mixed methods research and qualitative methodology as described by Hesse-Biber (2010)[8].

Explanatory sequential mixed methods study

Our research objectives were answered by means of an *explanatory sequential mixed-method* design, adopted in accordance to Hesse-Biber (2010). The study is "explanatory" as the focus is on understanding experiences in-depth, and "sequential" because quantitative and qualitative data strands were analysed independently and integrated at the phase of interpretation, allowing for a complementation and deeper understanding of both the quantitative and qualitative outcome findings. This design can be summarized as "quan → QUAL": a qualitative strand was added secondary to the quantitative outcome evaluation yet becoming the most important focus of the study. Particularly, the quantitative strand served to define the samples for the qualitative analysis (cf. first "descriptive layer", the interviews were analysed subsequently (cf. second and third "explorative" and "explanatory" layer), and in a

[8] "Qualitative methodologies should not be mistaken for qualitative methods. Qualitative methodologies, as noted earlier, use quantitative as well as qualitative methods. The same rule applies for quantitative methodologies. After all, the method is but the tool; the methodology determines the way in which the tool will be utilized." (Hesse-Biber, 2010; p.17). In general, Hesse-Biber argues to centralize the *research problem* when conducting mixed methods rather than the mixed methods *design.*

final stage, the relationship between patients' lived experiences and quantitative outcome classification were examined. Characteristic to a mixed methods study, the studies evolved iteratively; findings from one study led to additional questions or finetuning of questions for the following study, resulting in an incremental process of knowledge (or theory) building (Hesse-Biber, 2010). By going back and forth between the analyses of the samples, the quantitative and qualitative data, our understanding of outcome and change processes as found in patients' experiences grew progressively. An overview of this mixed methods, iterative process is depicted in Figure C and will be described when giving an overview of the conducted studies.

Primary outcome measure. For the purpose of this project we explicitly followed some of the standard procedures in standard psychotherapy outcome research in order to address our research objectives in a representative example of the type of findings that are relied on to inform policy decisions. Particularly, we relied on a primary outcome measure to evaluate the clinical significance of patients' outcome (see Atkins et al., 2005; De Los Reyes et al., 2011). We selected the *Beck Depression Inventory* (BDI-II; Beck, Steer, & Brown, 1996; van der Does, 2002), in accordance to the focus of the RCT this project is situated in (i.e., the efficacy of psychotherapy for major depression, cf. infra). The BDI is one of the most-often used measures of symptom severity in psychotherapy outcome research (cf. representative example; Hill & Lambert, 2004), shows sufficient psychometrical qualities as well as appropriate specificity and sensitivity to evaluate depression complaints (van der Does, 2002). For these reasons, the measure deemed more appropriate in comparison to the Hamilton Rating Scale for Depression (HRSD; Hamilton, 1967)[9]. Especially, the HRSD is an *observer*-rater screening

[9] With regard to major depression, clinical trials mostly rely on the Hamilton Rating Scale for Depression (HRSD; Hamilton, 1967) as primary outcome research. The discussion on how to define recovery and remission from depression by means of the HRSD, for instance the predefined cut-off, has a long-standing tradition in psychiatric literature, and we refer the reader to this interesting line of work (see Frank et al, 1991; Romera et al., 2011; Rush et al., 2006; Zimmerman et al., 2012). We will refer to this work at times to draw parallels between our findings and the broader discussion on how to evaluate significant changes in the treatment of major depression. Yet, by choosing the BDI it allows us to bring the discussion to the less-often addressed

instrument, often-used in clinical trials and particularly psychiatric literature. Given the focus on patients' perspectives, we explicitly chose to rely on a patient self-report measure. Also, different criteria and definitions to define recovery are used in this strand of psychiatric literature (see Zimmerman et al., 2012), that do not cover the notion of clinical significance as defined by Jacobson and Truax (1991). Moreover, the HRSD is criticized on its limited specificity and sensitivity to evaluate depression, indeed the HRSD focusses strongly on (for instance) anxiety and insomnia which is not in line with the dominant diagnostical description of depression (e.g., as described in the Diagnostical and Statistical Manual of Mental Disorder, DSM; APA, 2013). Moreover, given the focus of the RCT on the efficacy of psychotherapy for major depression, a construct-related measure was also preferred over less specific general wellbeing measures (e.g., symptom checklist inventory, SCL-90; Derogatis, 1992, or Outcome Questionnaire-45; Lambert et al., 1996).

Sampling strategy. Following our central research objective to explore patients' experiences explicitly *in relation to* the statistical outcome classifications, we applied a *purposeful sampling strategy* to select our research participants. Specifically, we sampled patients based on a "predetermined criterion of importance" (i.e., criterion sampling), as often used in mixed methods research projects (Palinkas et al., 2015). Generally, this allows to elaborate on one strand of the data (e.g., the quantitative outcome findings) by means of the other (e.g., the in-depth analysis of patients' experience), or to integrate and compare findings from both analyses which allows for a better understanding or gaining new insights (as described by Hesse-Biber, 2010; cf. Patton, 2002). Specifically, the samples were primarily identified based on patients' outcome category as defined using the JT method of clinical significance, resulting into four samples: recovered, improved, non-improved and deteriorated patients. Relying on the BDI-II (van der Does, 2002) as our outcome measure, this implied the

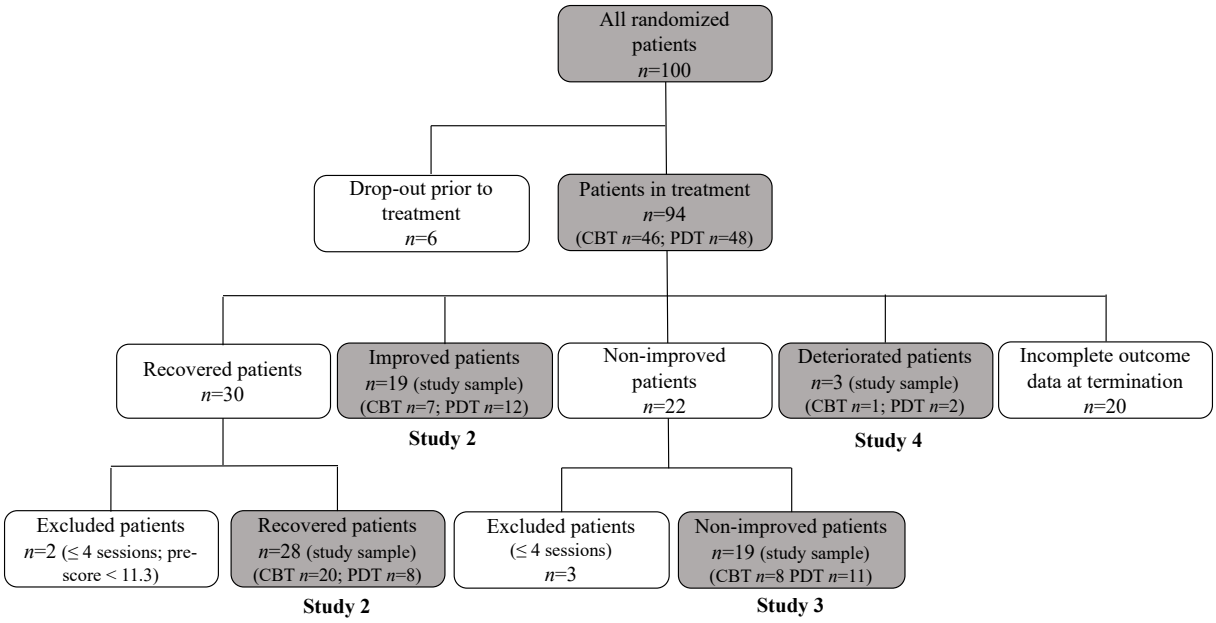
definition of clinical significance by Jacobson and Truax (1991) and particularly to the field of psychotherapy outcome research.

following: in order to reach reliable change on the Dutch BDI, a patient must show change in scores equal to or larger than 9.6. The clinical cut off is set at 11.3 (based on the internal consistency of 0.92, van der Does, 2002). Cut-off C was used to calculate the cut-off between the functional and dysfunctional population, as normative data are available for both patient (*M*=21.3; *SD*=12.2) and non-patient sample (*M*=6.2; *SD*=6.2) and both samples are overlapping (van der Does, 2002; see Lambert et al., 2008). While selecting the patients for the four samples from the larger RCT, we applied a variation of what is more generally known as "total population" strategy (i.e., selecting all patients in the RCT who share the primary criterion; the RCT is interpreted as our "population" in this context), as we aimed to investigate *all voices* represented within the pre-defined categories, rather than selecting a part of them (as for instance done by means of typical case or maximum variation sampling; cf. Patton, 2002). Some additional criteria were used to guarantee appropriate classification of patients according to the JT method. In order to be included for our study, patients needed a pre-score above the clinical cut-off of 11.3, in order to be able to reach clinically significant change; one patient was excluded based on this criterion[10]. Secondly, given our aim to study "psychotherapy outcome", we deliberately differentiated patients who "dropped out" from therapy *prematurely* (who can be assumed to be characterized by different experiences, see O'keeffe, Martin, Target & Midgley, 2019). Concretely, in line with dominant definitions of drop out applied in the research literature we excluded patients who ended therapy within four sessions (see Barrett et al., 2006). Four patients were excluded based on this criterion. This resulted into a recovered sample consisting of 28 patients, an improved sample with 19 patients, a non-improved sample with 19 patients and three cases who deteriorated. Figure B gives an overview of the sample selection applied in studies 2-4: seventy-four out of 100 randomized patients in the RCT

[10] This is a consequence of the chosen outcome measure in this study (i.e., the BDI) which differs from the chosen primary outcome measure in the Ghent Psychotherapy Study (i.e., the Hamilton Rating Scale for Depression). See primary outcome measure for an argumentation for choosing the BDI over the HRSD.

provided complete post-data and could thus be considered for our sample selection (we will further describe the samples when introducing the separate chapters).

Figure B. Overview of the samples selected for study 2-4



Note. Inclusion criteria: Patients with complete data post treatment (audiotaped interview and quantitative outcome scores on BDI). Exclusion criteria: Drop-out within four sessions of therapy, pre-score below the clinical cut-off for the BDI (11.3)

First-person perspective: giving voice to the patients' narrative experiences

In the applied sequential explanatory mixed method designs, the qualitative strand in this research project is most prominent. We focused on patients' perspectives in particular. The premise of this dissertation is that the patients' perspective is a valid perspective that can offer a valuable contribution to the understanding and measurement of outcome and therapy effects, especially when clinical meaningfulness is aimed for (Bracken et al., 2012; Strupp & Hadley, 1977). The term "patient perspective" or "first-person perspective" have, however, been used in various ways in research literature and it is therefore necessary to clarify how they should be interpreted in this dissertation. In each specific study (cf. infra), we will refer to "patient's view", "lived experience", "idiosyncratic perspective" or "subjective experience" to indicate

that we focus on the experiences as narrated by the patient, contrasting the researcher or therapist perspective (which can be considered a "third-person perspective"; Englebert, Follet & Valentiny, 2017). In this definition, self-report questionnaires containing predefined questions and answers are hence not considered a first-person perspective (Hill et al., 2013). Patients' perspectives were inquired by means of the Client Change Interview (CCI) as developed by Elliott, Slatick and Urman (2001), which comprises the domains of experienced changes during therapy, helping and hindering elements in and outside of therapy and has been designs as a semi-structured means to inquire patients' experiences.

Qualitative, data-driven analyses departing from a contextual perspective

Given the aim for in-depth exploration of patients' experienced changes and grasping the change processes or influential factors that help explain these experienced changes, we relied on data-driven qualitative methods that allowed for in-depth inquiry, mainly relying on Grounded Theory and Phenomenological principles (Ponterotto, 2005; Strauss & Corbin, 1990). These methods can be used to provide rich description and interpretation, with the aim to generate tentative conceptual models or theory, in line with our aim to deepen the understanding of outcome or change and the change processes leading up to it (Fassinger, 2005; Charmaz, 2014). Importantly, in contrast to approaches that integrate theoretical interpretations (e.g., interpretative phenomenological analysis, IPA; Smith, Flower & Larkin, 2009), we solely relied on *patients' perspectives*, building the conceptual models in a bottom-up manner. This means our research aim to elucidate patients' experienced changes and perspectives on important change processes followed an inductive or data-driven approach (rather than a deductive angle; Braun & Clarke, 2006). The broader frame of reference that guided the research and analyses can be described as a "contextual perspective", which departs from the assumption that patients' experiences and perspectives are influenced by the broader social context (Boyatzis, 1998). As such we depart from a constructivist-interpretative paradigm

(Braakmann, 2015), an epistemology that assumes the existence of "multiple subjective realities that consist of stories or meanings produced or constructed by individuals within their "natural" settings." (Hesse-Biber, 2010; p. 18).

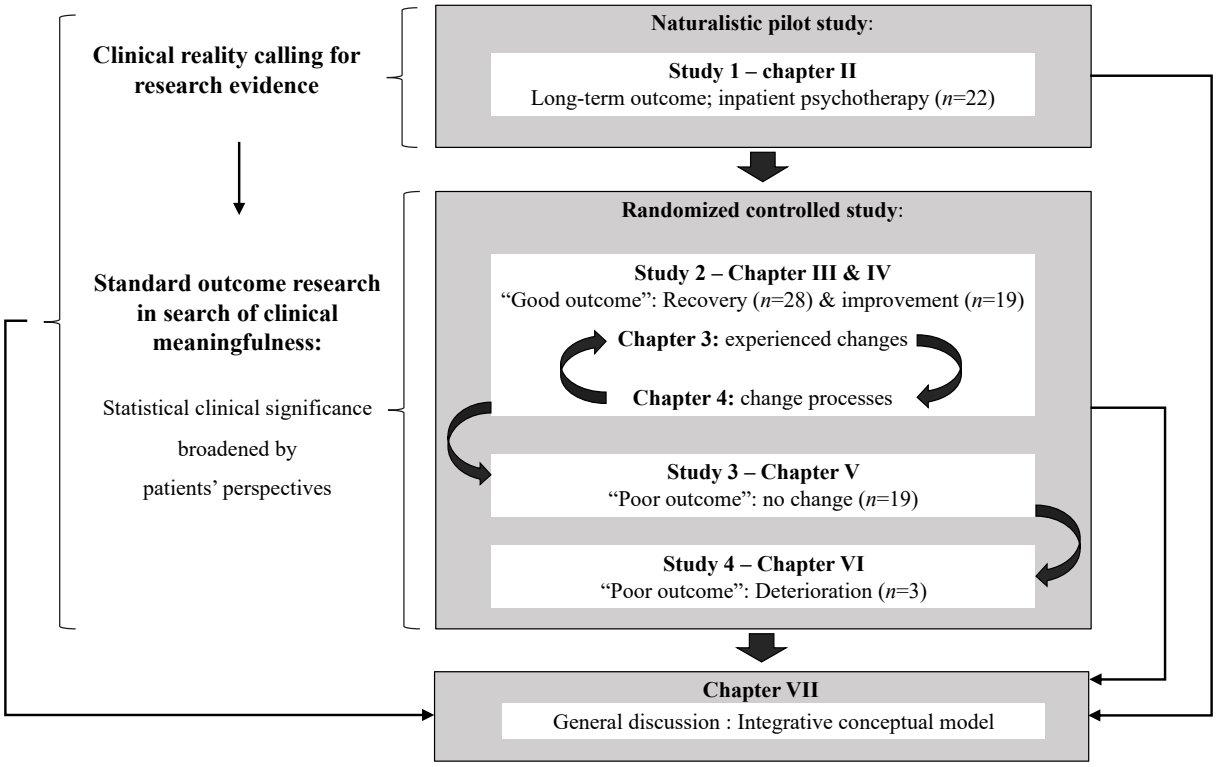
Situating the present dissertation within the process-outcome field: A focus on "change"

In scrutinizing patients' perspective on psychotherapy outcome, we make use of the term "change" as this enables to bridge the historical arbitrary divide between "outcome" and "process" in line with the recommendations by Kiesler (1966), Greenberg and Pinsof (1986), and Elliott (2010). This can be considered a "process approach" that views outcome as a fluid and continuous process that is not "best measured" at termination or any other single point. Given our focus on patients' experiences of change, this understanding aligns our research aim to elucidate what is meaningful to patients (inside or outside of the therapy session, and at any given time) beyond predefined indications of what outcome *is* (cf. Carey et al., 2006). Similarly, in studying patients' perspectives on change processes, we did not solely ask patients about the helping and hindering elements of psychotherapy but broadened the perspective of inquiry to other potentially influences over course of psychotherapy (i.e., inside and outside of the setting; cf. contextual perspective; McLeod, 2013).

Overview of the conducted studies and chapters

The objectives of this dissertation were answered by means of four mixed methods studies, all tackling several more specific research questions and topics. The findings of these studies will be presented in five consecutive chapters. Chapter II presents a naturalistic study on long-term outcome and serves as the pilot study for the main research project which takes place in the context of an RCT. The findings of this research are presented in chapter III to VI. In chapter VII, the findings from all separate mixed methods studies are integrated in an integrative conceptual model. Figure C provides a visual overview of the iterative mixed methods project and the consecutive studies and chapters.

Figure C. Overview of chapters and the conducted studies



Note. The arrows indicate the waves in which the studies were conducted, representing an iterative and progressive research process (Hesse-Biber, 2010). The naturalistic pilot study was conducted first and enabled rich exploration of patients' experienced changes and change processes. The findings of this study informed the studies that were subsequently conducted in context of the RCT. The general discussion (chapter VII) offers an integrative conceptual model based on all studies.

Pilot study: Clinical practice calling for research evidence. *Study 1* presented in *Chapter II* departs from a pressing clinical reality: treatment institutions must prove their effectiveness in order to be funded by government agencies or health care policies. Standardized pre-to-post research designs however are not adjusted to the complex clinical context of inpatient treatment settings, characterized by a combination of treatments within one setting and not reduceable to the sum of its parts. Moreover, this clinical reality also means that mere proof of effectiveness does not yet inform the institution and its clinicians on how and why their approach works. In order to improve their clinical care these are however crucial insights. Addressing these concerns within an inpatient setting in Flanders, Belgium, study 1 consisted of a mixed methods examination of former inpatients' long-term outcome. Twenty-

two former inpatients who engaged in one-year inpatient care were evaluated by means of the Outcome Questionnaire (Lambert et al., 1996) during and after treatment and interviewed in-depth five to six years after therapy. The following research questions were addressed by means of a data driven thematic analysis: 1) which changes have former inpatients experienced during and after inpatient psychotherapy, 2) which factors (in or outside treatment) they believe have influenced these changes and how can the explanatory factors help to understand individual differences between former patients' experienced changes? 3) how the perspectives of former inpatients complement findings of long-term outcome based on quantitative measurement. The study presents a tentative conceptual model of long-term outcome grounded in patients' experiences.

The naturalistic study provided rich in-depth interviews on patients' experiences, which informed the subsequent studies. The study served as a pilot study in this dissertation as it informed both the method and focus of the subsequent studies which were conducted within a randomized and standardized research setting. For instance, more interpretative qualitative methods (i.e., grounded theory and IPA) were selected in order to foster the construction of conceptual process models. Moreover, following from the observations in the naturalistic study, the implications of how outcome is conceptualized and measured in standard psychotherapy outcome research became a central research focus. Turning from a clinical reality in need of research evidence, we turned to a standardized research context in need of clinical relevance.

RCT: Standard outcome research in search of clinical meaningfulness. Addressing some of the key concerns in standard outcome research, a randomized controlled research context offered the ideal and unique setting to enable a systematic evaluation of the clinical meaningfulness of outcome. Concretely, mixed method data were gathered in the context of the Ghent Psychotherapy Study, an RCT on the treatment of major depression[11]. The research

[11] See protocol for more details on the RCT (Meganck et al., 2017). See https://www.ncbi.nlm.nih.gov/pubmed/28292331. The trial is registered at open science framework: https://osf.io/mf2d7/

questions addressed in this research project are not limited to the case of major depression, yet in study 2-4 they will be answered in regard to this representative case (Rottenberg et al., 2018). As major depression is one of the most prevalent mental health problems (WHO, 2017) it has been the object of study in numerous efficacy studies (Cuijpers et al., 2014[12]), thus representing both a topical and representative case to address our research objectives. The implications for the study and understanding of major depression will be addressed throughout the chapters and we will return to this in the general discussion (chapter VII).

Following a mixed-methods philosophy, the studies were conducted in an iterative, back-and-forth manner (see Figure C). First, we present the study conducted on patients' showing "good outcome" (i.e., recovered and improved patients), thereafter we turn to the studies on those patients who were marked as showing "negative outcome" (i.e., non-improved and deteriorated patients).

"Good outcome". In *chapter III* and *chapter IV,* we examine the positive end of the outcome continuum, looking into the experiences of those patients showing "good" psychotherapy outcome. Based on the JT definition of clinical significance (Jacobson & Truax, 1991), good outcome can be understood as those patients showing recovery (i.e., clinically significant change) and those showing improvement (i.e., reliable change). While recovery is typically considered the gold standard of psychotherapy outcome, also improvement is considered an indication of successful psychotherapy outcome and generally reported as such (e.g. McElvaney & Timulak, 2013). The meaningfulness of the differentiation between recovery and improvement in light of patients' experiences however remains underexplored. Therefore, both recovered (n=28) and improved (n=19) patients were included in the study. The research questions addressed in this study are: 1) how do patients who recovered and improved

[12] See database of randomized controlled and comparative studies on the effects of psychotherapy for adult depression, developed by Pim Cuijpers and colleagues: http://www.evidencebasedpsychotherapies.org

in depression symptomatology experience their outcome (i.e., patients' general outcome appraisal, e.g., do they feel improved themselves)? 2) what changes have these patients experienced since the start of therapy? 3) what do patients, who show recovered and improved outcome, report as contributions to change? 4) what differences can be observed between experiences of patients classified as either recovered or improved? As depicted in Figure C, the study developed in an iterative manner and various ways to represent the findings were applied before arriving at the final conceptual models (this is described in more detail in chapter III and IV). In order to enable considerable elaboration on all research questions addressed, the study was eventually presented in two separate chapters, in order to do justice to the richness of the data in the study.

Chapter III covers the first two questions, hence providing a thorough exploration of outcome from a first-person perspective (cf. exploratory layer; multi-layered analysis of change). Specifically, we examine what "good outcome" means in light of patients' perspectives. As such, we particularly address the methodological concern in the field of psychotherapy as to how "good outcome" should be studied and conceptualized and we investigate whether the JT classification differentiates recovered and improved patients' outcome in a clinically meaningful way. Specifically, we explore the experienced changes of patients who can be categorized as "recovered" and "improved" based on self-reported depression symptoms in order to gain insight into what lays within and beyond the header of "good" psychotherapy outcome.

Chapter IV looks into the third and fourth question, presenting an elaborate study of change mechanisms from a first-person perspective (i.e., explanatory layer). Specially, we build onto recovered and improved patients' experienced changes and elucidate the change processes underlying "good outcome" as perceived by these patients. In the general discussion (chapter

VII), we will offer an integrative discussion of the findings, synthesizing the experienced changes ("exploratory layer") and change processes ("explanatory layer").

"Poor outcome". In *chapter V* and *chapter VI,* we expand the investigation of how outcome should be studied and conceptualized on the negative side of the outcome continuum. "Poor" outcome is typically understood as those patients showing no reliable change or non-response to treatment and patients showing deterioration or worsening in complaints.

In *chapter V*, we investigate the perspectives of the patients who did not show reliable improvement in self-reported symptom severity (n=19) in order to explore how non-improvement in pre-to-post symptom severity can be understood in relation to the experience of depressed patients themselves. In this study we translate the question regarding change and process mechanisms to the situation of a possible absence of change. In doing so, we explicitly assumed that patients who show no reliable change in outcome scores, can potentially show negative, positive or no alterations in experiences. The choice to study non-improved patients separately from those who deteriorated in therapy is motivated by the assumption that lack of improvement and worsening must be regarded as different phenomena, influenced by different mechanisms (Lambert, 2011).

In *chapter VI* we therefore provide an in-depth examination of the three cases in the RCT sample who showed deteriorated outcome scores pre-to-post therapy on the BDI. Using an Interpretative Phenomenological Analytic approach (IPA; Smith et al., 2009) we explore the relationship between outcome measures and patients' experiences in the case of poor outcome in more detail. Moreover, studying three cases allowed the in-depth exploration of the underexplored phenomenon of worsening in psychotherapy. A central explorative research question was accompanied by several secondary more specific research questions, as typical for IPA methodology (Smith et al., 2009). The latter were formulated building onto the first-person knowledge gathered in the previous studies (i.e., study 1, 2 and 3). Primarily, we

explored how patients who show deterioration at the level of symptom severity describe their outcome and therapy process. Secondly, as the findings from the previous mixed methods studies identified the role of the patient, therapist, therapy and external context, for each deteriorated case we asked specific questions about how these factors were represented in their narratives.

Chapter VII offers a thorough discussion of the findings, presenting an integrative conceptual model of outcome from patients' perspectives based on the findings gathered in all preceding waves of the mixed methods study (Hesse-Biber, 2010). In a final step the comparison of the findings of all mixed methods studies, allow us to draw some general conclusions on the clinical value of the statistical definition of clinical meaningfulness in light of patients' experiences.

References

American Psychological Association (APA). (2006). Evidence-based practice in psychology. American Psychologist, 61(4), 271–285.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (5th ed.).* Arlington, VA: American Psychiatric Association Publishing.

Atkins, D. C., Bedics, J. D., McGlinchey, J. B., & Beauchaine, T. P. (2005). Assessing clinical significance: Does it matter which method we use? *Journal of Consulting and Clinical Psychology*, *73*(5), 982–989. https://doi.org/10.1037/0022-006X.73.5.982

Barrett, M. S., Chua, W.-J., Crits-Cristoph, P., Gibbons, M. B., Casiano, D., & Thompson, D. (2008). Early withdrawal from mental health treatment: implications for psychotherapy practice. *Psychotherapy (Chicago Ill), 45*(2), 247–267. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2762228/

Blanton, H., & Jaccard, J. (2006). Arbitrary metrics in psychology. *The American Psychologist*, *61*(1), 27–41. doi: 10.1037/0003-066X.61.1.62

Bracken, P. et al. (2012). Psychiatry beyond the current paradigm. *British Journal of Psychiatry, 201,* 430-434.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3,* 77-101. doi: 10.1191/1478088706qp063oa

Beck A.T., Steer R.A., & Brown G.K. (1996). *Manual for the Beck Depression Inventory II.* San Antonio, TX: Psychological Corporation

Boyatzis, R. E. (1998). *Transforming qualitative information: thematic analysis and code development.* London, England: Sage

Braakmann, D. (2015). Historical paths in psychotherapy research. In O. C. G. Gelo, A. Pritz & B. Rieken (Eds.), *Psychotherapy research: foundations, process and outcome* (pp. 39-66). Vienna, Austria: Springer.

Carey, T. A., Carey, M., Mullan, R. J., Murray, L. K., & Spratt, M. B. (2006). Psychological Change: What changes and how does it occur? A critical review. *Counselling Psychology Review, 21*(4) 28-38

Carey, T. A., Carey, M., Stalker, K., Mullan, R. J., Murray, L. K., & Spratt, M. B. (2007). Psychological change from the inside looking out: A qualitative investigation. *Counselling and Psychotherapy*

Castonguay, L., Barkham, M., Lutz, W., & McAleavey, A. (2013). Practice-oriented research: Approaches and applications. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior*

Charmaz, K. (2014). *Constructing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage.

Consort (2010). CONSORT 2010 checklist of information to include when reporting a randomised trial. Retrieved from www.consort-statement.org.

Cuijpers, P., Karyotaki, E., Weitz, E., Andersson, G., Hollon, S. D., & van Straten, A. (2014). The effects of psychotherapies for major depression in adults on remission, recovery and improvement: A meta-analysis. *Journal of Affective Disorders*, *159*, 118–126. https://doi.org/10.1016/j.jad.2014.02.026

Cuijpers, P., Reijnders, M., & Huibers, M. J. H. (2019). The Role of Common Factors in Psychotherapy Outcomes. *Annual Review of Clinical Psychology, 15,* 207-231

Dattilio, F. M., Edwards, D. J. A., & Fishman, D. B. (2010). Case studies within a mixed methods paradigm: Toward a resolution of the alienation between researcher and practitioner in psychotherapy research. *Psychotherapy: Theory, Research, Practice, Training*, *47*(4), 427–441. https://doi.org/10.1037/a0021181

De Los Reyes, A., Kundey, S. M. A., & Wang, M. (2011). The end of the primary outcome measure: A research agenda for constructing its replacement. *Clinical Psychology Review, 31*(5), 829–838. https://doi.org/10.1016/j.cpr.2011.03.011

Derogatis, L.R. (1992). *SCL-90-R Administration, scoring and procedures manual (2nded). T*owson, MD: Clin Psychom Res Inc.

Englebert, J., Follet, V., Valentiny, C. (2017). Anorexia Nervosa and first-person perspective: altruism, family system and body experience. *Psychopathology, 51*, 24-30. DOI: 10.1159/0004852629

Elliott, R. (2010) Psychotherapy change process research: Realizing the promise. *Psychotherapy Research, 20*(2), 123-135, DOI: 10.1080/10503300903470743

Elliott, R., Slatick, E., & Urman, M. (2001). Qualitative change process research on psychotherapy: Alternative strategies. In J. Frommer & D. L. Rennie (Eds.), *Qualitative psychotherapy research: Methods and methodology* (pp. 69–111). Lengerich, Germany: Pabst Science.

Evans, I. M., & Fletcher, A. (2013). Focus on change: implications for the understanding and evaluation of psychological interventions. *Psychologia, 56*(2), 113–130. https://doi.org/10.2117/psysoc.2013.113

Fassinger, R. E. (2005). Paradigms, praxis, problems, and promise: Grounded theory in counseling psychology research. *Journal of Counseling Psychology, 52*, 156– 166. doi:10.1037/0022-0167.52.2.156

Frank et al. (1991). Conceptualization and rationale for consensus definitions in terms of major depressive disorder: remission, recovery, relapse, and recurrence. *Archives of General Psychiatry, 48*, p.851-855

Goldfried, M. R., & Wolfe, B. E. (1996). Psychotherapy practice and research. Repairing a strained alliance. *The American Psychologist*, *51*(10), 1007–1016.

Greenberg, L. S. (1991). Research on the process of change. *Psychotherapy Research*, *1*(1), 3–16.

Greenberg, L. S. (1999). Ideal psychotherapy research: A study of significant change processes. *Journal of Clinical Psychology*, *55*(12), 1467–1480.

Greenberg, L. & Pinsof, W.M. (1986). Process research: current trends and future perspectives. In L. Greenberg and W. M. Pinsof (Eds) *The psychotherapeutic process. A research handbook* (pp.3-20). New York: NY, The Guilford Press

Hamilton, M. (1967). Development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology, 1967*(6), 278–96. https://doi.org/10.1111/j.2044-8260.1967.tb00530.x

Hesse-Biber, S. (2015). Mixed Methods Research: The "Thing-ness" Problem. *Qualitative Health Research*, *25*(6), 775–788. https://doi.org/10.1177/1049732315580558

Hill, C. E., & Lambert, M. J. (2004). Methodological issues in studying psychotherapy processes and outcomes. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behaviour change* (5th Ed.) (pp. 84-135) Hoboken, New Jersey: John Wiley & Sons, Inc

Hill, C. E., Chui, H., & Baumann, E. (2013). Revisiting and reenvisioning the outcome problem in psychotherapy: An argument to include individualized and qualitative measurement. *Psychotherapy, 50*(1), 68–76

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

Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy*, *15*(4), 336–352. https://doi.org/10.1016/S0005-7894(84)80002-7

Jadad, A., & Enkin, M. (2007). Randomized controlled trials: Questions, answers, and musings (2nd ed.). Malden, MA: Blackwell Publishing.

Kazdin, A. E. (1977). Assessing the clinical or applied importance of behavior change through social validation. *Behavior modification*, *1*(4), 427-452.

Kazdin, A. E. (1999). The meanings and measurement of clinical significance. *Journal of Consulting and Clinical Psychology*, 332–339.

Kazdin, A. E. (2001a). Almost clinically significant (p<.10): Current measures may only approach clinical significance. *Clinical psychology: Science and practice,* 8(4), 455-462. https://onlinelibrary.wiley.com /doi/pdf/10.1093/clipsy.8.4.455

Kazdin, A. E. (2001b). Bridging the Enormous Gaps of Theory With Therapy Research and Practice. *Journal of Clinical Child & Adolescent Psychology*, *30*(1), 59–66. https://doi.org/10.1207/S15374424JCCP3001_7

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

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

Kiesler, D. J. (1966). Some myths of psychotherapy research and the search for a paradigm. *Psychological Bulletin, 65*(2), 110–136.

Kiesler, D. J. (1971). Patient experiencing level and successful outcome in individual psychotherapy of schizophrenics and psychoneurotics. J*ournal of consulting and clinical psychology, 37*, 370-385

Kiesler, D. J. (1973). *The process of psychotherapy.* Chicago: Aldine

Lambert, M. J., Hansen, N. B., Umphress, V., Lunnen, K., Okiishi, J., Burlingame, G. M., & Reisinger, C. W. (1996). *Administration and scoring manual for the OQ-45.2*. East Setauket, NY: American Professional Credentialing Services.

Lambert, M. J. (2013). The efficacy and effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behaviour change* (6th Ed.) (pp. 169-218). Hoboken, New Jersey: John Wiley & Sons, Inc.

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. DOI:10.1080/10503300902849483

Lambert, M. J., Hansen, N. B., & Bauer, S. (2008). Assessing the Clinical Significance of Outcome Results. In A. M. Nezu & C. M. Nezu (Eds.). *Evidence-based outcome research. A Practical Guide to Conducting Randomized Controlled Trials for Psychosocial Interventions* (pp. 359-378). New York, NY: Oxford University Press

Levitt, H. M., Stanley, C. M., Frankel, Z., & Raina, K. (2005). An Evaluation of Outcome Measures Used in Humanistic Psychotherapy Research: Using Thermometers to Weigh Oranges. *The Humanistic Psychologist*, *33*(2), 113–130. https://doi.org/10.1207/s15473333thp3302_3

Levitt, H. M., Pomerville, A., & Surace, F. I. (2016). A qualitative meta-analysis examining clients' experiences of psychotherapy: A new agenda. *Psychological Bulletin*, *142*(8), 801–830.

Maassen, G. H. (2000). Principles of Defining Reliable Change Indices. *Journal of Clinical and Experimental Neuropsychology*, *22*(5), 622–632. https://doi.org/10.1076/1380-3395(200010)22:5;1-9;FT622

McElvaney, J., & Timulak, L. (2013). Clients' experience of therapy and its outcomes in 'good' and 'poor' outcome psychological therapy in a primary care setting: An exploratory study. *Counselling and Psychotherapy Research*, *13*(4), 246–253. https://doi.org/10.1080/14733145.2012.761258

McGlinchey, J. B., Atkins, D. C., & Jacobson, N. S. (2002). Clinical significance methods: Which one to use and how useful are they? *Behavior Therapy*, *33*(4), 529–550. https://doi.org/10.1016/S0005-7894(02)80015-6

McLeod, J. (2011). Qualitative research in counselling and psychotherapy. London, England: Sage.

McLeod, J. (2013). Qualitative research. Methods and contributions. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed.) (pp. 49-84). New York, NY: John Wiley & Sons.

McLeod, J. (2016). *Using research in counselling and psychotherapy.* London, England: sage

Meganck, R., et al. (2017). The Ghent Psychotherapy Study (GPS) on the differential efficacy of supportive-expressive and cognitive behavioral interventions in dependent and self-critical depressive patients: study protocol for a randomized controlled trial. *Trials, 18*(126), 1-11. DOI:10.1186/s13063-017-1867-x

Midgley, N., Target, M., & Smith, J. (2006). The outcome of child psychoanalysis from the patient's point of view: a qualitative analysis of a long-term follow-up study. *Psychology and Psychotherapy*, *79*, 257–269. doi: 10.1348/147608305X52694

Midgley, N., Ansaldo, F., & Target, M. (2014). The meaningful assessment of therapy outcomes: Incorporating a qualitative study into a randomized controlled trial evaluating the treatment of adolescent depression. *Psychotherapy, 51*(1), 128-137. Doi:10.1037/a0034179

Morley, S., Williams, A. & Eccleston, C. (2013). Examining the evidence about psychological treatments for chronic pain: Time for a paradigm shift? *Pain, 154*(10), 1929–1931. doi: 10.1016/j.pain.2013.05.049.

Munder, T., Flückiger, C., Leichsenring, F., Abbass, A. A., Hilsenroth, M. J., Luyten, P., ... & Wampold, B. E. (2019). Let's Distinguish Relative and Absolute Efficacy to Move Psychotherapy Research Forward. Zeitschrift für Psychosomatische Medizin und Psychotherapie, 65(2), 178-182.

Ogles, B. M., Lunnen, K. M., & Bonesteel, K. (2001). Clinical significance: History, application, and current practice. *Clinical Psychology Review, 21*(3), 421-446. https://doi.org/10.1016/S0272-7358(99)00058-6

Ogles (2013). Measuring change in psychotherapy research. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behaviour change* (6th Ed.) (pp. 134-166). Hoboken, New Jersey: John Wiley & Sons, Inc.

O'Keeffe, S., Martin, P., Target, M., & Midgley, N. (2019) 'I Just Stopped Going': A Mixed Methods Investigation Into Types of Therapy Dropout in Adolescents With Depression. *Frontiers in Psychology. 10*(75). doi: 10.3389/fpsyg.2019.00075

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful Sampling for Qualitative Data Collection and Analysis in Mixed Method Implementation Research. *Administration and Policy in Mental Health and Mental Health Services Research*, *42*(5), 533–544. https://doi.org/10.1007/s10488-013-0528-y

Patton, M. (2002) *Qualitative Research and Evaluation Methods* (3rd Ed). Thousand Oaks, CA: Sage.

Ponterotto, J. G. (2005). Qualitative research in counseling psychology: A primer on research paradigms and philosophy of science. *Journal of Counseling Psychology*, *52*(2), 126–136. https://doi.org/10.1037/0022-0167.52.2.126

Rodgers, B. (2003) An exploration into the client at the heart of therapy: a qualitative perspective. *Person-Centered and Experiential Psychotherapies 2*(21), 19–30. doi:10.1080/14779757.2003.9688290

Rodgers, B., & Elliott, R. (2013). Qualitative methods in psychotherapy outcome research. In O. C. G. Gelo, A. Pritz & B. Rieken (Eds.), *Psychotherapy research: foundations, process and outcome* (pp. 559-579). Vienna, Austria: Springer.

Romera, I., Pérez, V., Menchón, J. M., Polavieja, P., & Gilaberte, I. (2011). Optimal cutoff point of the Hamilton Rating Scale for Depression according to normal levels of social and occupational functioning. *Psychiatry Research*, *186*(1), 133–137. https://doi.org/10.1016/j.psychres.2010.06.023

Rottenberg, J., Devendorf, A. R., Kashdan, T. B., & Disabato, D. J. (2018). The Curious Neglect of High Functioning After Psychopathology: The Case of Depression. *Perspectives on Psychological Science*, *13*(5), 549–566. https://doi.org/10.1177/1745691618769868

Rush, A. J., Kraemer, H. C., Sackeim, H. A., Fava, M., Trivedi, M. H., Frank, E., … Schatzberg, A. F. (2006). Report by the ACNP Task Force on Response and Remission in Major Depressive Disorder. *Neuropsychopharmacology*, *31*(9), 1841–1853. https://doi.org/10.1038/sj.npp.1301131

Slife, B. D. (2004). Theoretical challenges to therapy practice and research: the constraint of naturalism. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behaviour change* (5th Ed.) (pp. 44-83) Hoboken, New Jersey: John Wiley & Sons, Inc

Smith, J.A., Flower, P., & Larkin, M. (2009). *Interpretative Phenomenological Analysis: Theory, Method and Research*. London: Sage.

Strauss, A., & Corbin, J. (1990). *Basics of qualitative research techniques and procedures for developing grounded theory*. London: Sage Publications.

Strupp, H. H., & Hadley, S. W. (1977). A tripartite model of mental health and therapeutic outcome: With specific reference to negative effects in psychotherapy. *American Psychologist, 32,* 187–196. doi:10.1037/0003- 066X.32.3.187

Tashakkori, A., & Teddlie, C. (Eds.). (2003). *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage.

Timulak, L. (2007). Identifying core categories of client-identified impact of helpful events in psychotherapy: A qualitative meta-analysis. *Psychotherapy Research*, *17*, 310–320. https://doi.org/10.1080/10503300600608116

Timulak, L. (2010). Significant events in psychotherapy: An update of research findings. *Psychology and Psychotherapy: Theory, Research and Practice*, *83*(4), 421–447. https://doi.org/10.1348/147608310X499404

van der Does, A.J.W. (2002). *BDI-II-NL. Handleiding. De Nederlandse versie van de Beck Depression Inventory-2nd edition.* Lisse: Harcourt Test Publishers.

Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work*. New York: Routledge.

Westen, D., Novotny, C. M., & Thompson-Brenner, H. (2004). The Empirical Status of Empirically Supported Psychotherapies: Assumptions, Findings, and Reporting in Controlled Clinical Trials. *Psychological Bulletin, 130*(4), 631–663. doi: 10.1037/0033-2909.130.4.631

Wise, E. A. (2004). Methods for Analyzing Psychotherapy Outcomes: A Review of Clinical Significance, Reliable Change, and Recommendations for Future Directions. *Journal of Personality Assessment*, *82*(1), 50–59. https://doi.org/10.1207/s15327752jpa8201_10

World Health Organization (2017). *Depression and other common mental disorders: Global health estimates*. Retrieved from http://apps.who.int/iris/bitstream/10665/254610/1/WHO-MSD-MER-2017.2

Zimmerman, M., Martinez, J. A., Attiullah, N., Friedman, M., Toba, C., Boerescu, D. A., & Rahgeb, M. (2012). Why do some depressed outpatients who are in remission according to the Hamilton Depression Rating Scale not consider themselves to be in remission? *Journal of Clinical Psychiatry, 73*(6), 790-795. https://doi.org/10.1002/da.21987

II. Understanding Long-term Outcome from the Patients' Perspective: A Mixed Method Naturalistic Study on Inpatient Psychotherapy

Abstract

The complex phenomenon of psychotherapy outcome requires further conceptual and methodological developments that facilitate clinically meaningful research findings. In this study, we rely on an idiosyncratic and process-oriented understanding of treatment effects in order to investigate long-term outcome. The latter follows several critical concerns about the arbitrary distinction between the process and outcome of psychotherapy (Hill & Lambert, 2004; Kiesler, 1966; Stiles, Shapiro & Harper, 1994) and recommendations for studying long-term outcome in particular (Ekroll & Holgorsen, 2015). A mixed methods naturalistic study was conducted in an inpatient psychotherapy setting. Long-term quantitative outcome data are complemented with a data-driven thematic analysis of interviews with 22 participants, five to six years after ending inpatient psychotherapy. Long-term outcome findings show improved well-being for the majority of former patients and this until five to six years after treatment. A conceptual model of long-term outcome is presented that comprises both a taxonomy of change and explanatory factors. From the patients' perspectives, long-term changes can be situated on different interrelated existential levels: reconnection to others and (the meaning of) life, a revelation, an altered self, life changes, and altered expectations and ideas about recovery and treatment. The complex interplay of the person, the therapy centre, the outside world and the evolution over time helped explain the experienced changes and individual differences. The findings support the value of an idiosyncratic and process-oriented understanding of outcome and recovery as well as substantiate the importance of multiple methods and perspectives when studying the effects of psychotherapy[13].

[13] De Smet, M., & Meganck, R. (2018). Understanding Long-term Outcome from the Patients' Perspective: A Mixed Methods Naturalistic Study on Inpatient Psychotherapy. *Psychologica Belgica, 58*(1), pp. 276–296. DOI: https://doi.org/10.5334/pb.432

Introduction

Understanding the effects or outcome of psychotherapy is essential for both research and clinical practice, yet the on-going discrepancy between both fields impedes a mutual dialogue (Castonguay, Barkham, Lutz & McAleavey, 2013; Kazdin, 2008). As outcome research has a huge impact on clinical practice through health care policy (APA, 2006; McLeod, 2016), the question of how to define and determine the effects of therapy remains a central issue. Evidence-Based Practice emphasises 'the *integration* of the best available research with clinical expertise in the context of patient characteristics, culture, and preferences' (APA, 2006, p. 273; emphasis added), yet in terms of policy this is often limited to the "best available evidence", which in turn is predominantly reduced to highly standardized efficacy studies. Unsurprisingly, given the complexity of clinical practice in general and treatment outcome in particular, this tendency is heavily criticized in both the academic literature and the clinical field (Gaudiano & Miller, 2013; Hill & Lambert, 2004; Westen, Novotny & Thompson-Brenner, 2004). To say the least, further conceptual and methodological developments are required in order to curtail the divide between research and clinical practice (Kazdin, 2008; McLeod, 2011).

In studying the effects of therapy, outcome has traditionally been separated from the process of treatment (Hill & Lambert, 2004), resulting in an arbitrary differentiation. This misconception has 'clouded the thinking regarding research design', especially in outcome research, as 'to some extent process research is outcome research, and outcome research is equivalent to process investigation' (Kiesler, 1966, p. 126). A unidirectional relationship between process and outcome is furthermore questioned (Stiles & Shapiro, 1989): 'process measurements taken from on-going therapy may be the result [outcome] rather than the cause [process] of clients' improvement' (Stiles, Shapiro & Harper, 1994, p. 42). Long-term follow-up studies have shown the value of investigating treatment outcome over a longer period of time, looking at treatment effects in a process-oriented manner, yet to date such long-term outcome studies remain limited (Ekroll & Rønnestad, 2016).

The impact and consequences of a demarcated or narrowed conception of outcome are most apparent in the so-called gold standard research design, the randomized controlled trial (RCT). High levels of standardization (e.g., strict in- and exclusion criteria, treatment protocols; Westen et al., 2004), an over-reliance on symptom-oriented outcome measures (Braakmann, 2015; Levitt, Stanley, Frankel & Raina, 2005), and the selection of one primary outcome measure to determine treatment outcome (De Los Reyes, Kundey & Wang, 2011), clearly simplify the understanding of treatment effects. Moreover, changes in outcome scores may not represent or correspond to an individual's functioning in daily life, nor can it give insight into how an individual interpreted the questions asked. Standardized outcome measures are therefore considered "arbitrary metrics" requiring real-life contextualization (Blanton & Jaccard, 2006; Hill, Chui & Bauman, 2013; Kazdin, 2001a; 2006).

Ergo, a growing interest in qualitative and mixed methods research emerges, as they allow for the contextualization of research findings and in-depth exploration of complex phenomena (Binder et al., 2016; Kazdin, 2008; Klein & Elliott, 2006; Midgley, Ansaldo & Target, 2014). In this regard, the perspective of the patient undergoing treatment is increasingly considered an essential source of information (Binder, Holgorsen & Nielsen, 2009; McLeod, 2016; Strupp & Hadley, 1977; Valkonen, Hanninen & Lindfors, 2011). This *first*-person perspective stands for the subjective experience, personal view or opinion of patients and must be considered distinct from a researcher or *third*-person perspective represented in the pre-defined questions and answers of standard outcome scales (Englebert, Follet & Valentiny, 2017; Englebert et al., 2018; Hill et al., 2013).

Opposed to standard outcome research, qualitative inquiry of patient perspectives has revealed important dimensions of change other than symptomatic changes (McLeod, 2011), for instance pointing at existential and interpersonal domains (cf. Binder et al., 2009; Nilsson, Svensson, Sandell & Clinton, 2007). Important individual differences in the experience of

outcome have also been observed depending on patients' personal background (Kühnlein, 1999; Valkonen et al., 2011) and the type of therapy (Nilsson et al., 2007). Moreover, it seems that the effects of psychotherapy must be understood in relation to the broader life span of patients (Midgley, Target & Smith, 2006) and other life domains (Binder et al., 2009). Together these studies show the relevance of investigating the patients' perspective on psychotherapy and outcome, the importance of a broader conception of treatment effects, and the value of understanding outcome within an idiosyncratic context.

The current study wishes to contribute to these lines of research within the particular context of inpatient psychotherapy, a setting that challenges the field with an even higher level of complexity due to its particular characteristics. The different therapy groups in inpatient treatment can create multiple therapeutic climates (Tschuschke & Dies, 1994), and various elements outside the therapy sessions (e.g., living together on a ward) can facilitate change (Leszcz, Yalom, & Norman, 1985). Hence, inpatient care cannot be considered equivalent to the sum of its parts. Additionally, inpatient populations are typically more complex, as they often suffer from more chronic mental difficulties and/or are situated within more complex social contexts compared to patients in outpatient psychotherapy. Due to these specific characteristics, the setting does not lend itself well for controlled research and has mostly been investigated by means of basic pre-post measurements (Kösters, Burlingame, Nachtigall & Strauss, 2006). Consequently, little is known about the particular effects of inpatient treatment and change processes over time. Using multiple methods and inquiring about patients' perspectives, can therefore offer a major contribution to the investigation of this specific and complex setting. In line with current ideas on practice-based evidence (Castonguay et al., 2013), we further believe naturalistic bottom-up inquiry offers important deepening and real-life knowledge, benefitting both research and clinical practice. Ultimately, practice-based research has the potential to circumvent the difficult translation of the context of research into the context

of clinical practice; at the same time addressing the sometimes-pressing influence of health care policy (cf. managed care) on the organisation of health care (Gaudiano & Miller, 2013).

In the context of inpatient treatment, the current study aims at understanding long-term outcome within the broader, personal life context of patients. Following a process-oriented understanding of outcome, we investigate which changes former inpatients have experienced during and after inpatient psychotherapy (RQ1) and which factors (in or outside treatment) they believe have influenced these changes (RQ2). Finally, we examine how the perspective of former inpatients complements findings of long-term outcome based on quantitative measurement (RQ3). To answer these objectives, a mixed methods design is deployed, incorporating quantitative long-term outcome findings and qualitative inquiry of patients' perspectives on changes over time. In an attempt to facilitate the implementation of findings into clinical practice, the present study takes place in a naturalistic setting.

Method

Setting

The study relies on naturalistic data gathered in an inpatient psychotherapeutic centre in Antwerp, Flanders (Belgium). The therapy centre offers intense, long-term inpatient psychotherapy to individuals with chronic mental difficulties, predominantly in the range of anxiety and mood disorders, and often in the context of attachment issues and personality disorders (cluster C). The overall approach of the treatment centre requires a certain level of stability, autonomy and self-reliance of patients in order to engage in the intense psychotherapeutic programme. For this reason, the centre must be considered distinct from a classical psychiatric ward; people suffering from acute psychological problems are referred.

After two initial intake interviews, a person engages in an exploratory treatment period of four weeks during which all forms of therapy are explored; only a minority is referred after this phase, when treatment is deemed to be ill-suited. During treatment, each patient receives

an individualized programme consisting of four to five forms of therapy, provided in six to nine sessions of individual or group therapy per week. Therapy is based on different frameworks and provided by trained and specialized therapists in cognitive behavioural, systemic and psychoanalytic therapy. Additionally, creative therapy is provided in the form of dance, movement, drama, visual and music therapy. Regular meetings with a psychiatrist are held regarding drug treatment; a personal mentor is the patient's contact throughout the treatment process. Finally, considerable free time is available for sports and other creative activities, all taking place in a quiet and natural surrounding environment.

Participants

Between September 2009 and August 2010, all individuals entering the treatment centre were asked to participate in an internal outcome study. At consecutive time points (i.e., at intake, every two months of treatment, post treatment, one-year and five-year follow-up) outcome was evaluated using the Outcome Questionnaire-45 (OQ-45, Lambert et al., 1996; de Jong et al., 2007). Forty-three participants entered the study; nine dropped out of the study (i.e., no post and follow-up data available). The remaining group (n = 34) provided data with variable response rate: 27 respondents at the end of treatment and 29 at one-year follow-up. The qualitative strand of this study focuses on the second follow-up inquiry, conducted five to six years after treatment termination. For this study, 27 participants (with nearly complete data) were contacted, 24 responded to our call for participation and 22 were willing to take part in the follow-up interview.

The research sample consisted of 19 women and 3 men, ranging in age from 31 - 61 (M = 47.95; SD = 8.94); all were of Caucasian European decent. Nine participants completed secondary education, 12 had degrees in higher education; seven of the participants were married, three were living together with their partners, and 11 participants indicated they were single (of which six were divorced) at the time of the interview (one participant did not mention

her marital status). Eleven of the participants were not working at the time, five of them were retired (often due to medical reasons), were registered as disabled, and one was unemployed at that time. The rest of the participants were employed, of which one was self-employed; three did voluntary work.

Patients were diagnosed at the start of treatment according to DSM-IV-TR (American Psychiatric Association, 2000). The recorded Axis I and Axis II diagnoses were: depressive disorder (n = 14), somatization disorder (n = 5), relational problems (n = 4), anxiety disorder (n = 1), PTSD (n = 2), adjustment disorder (n = 5), personality disorder NOS (n = 12), dependent PD (n = 4), avoidant PD (n = 4), obsessive compulsive PD (n = 3) and borderline PD (n = 1). The majority (n = 18) stayed for the maximum period of 12 months in treatment, one person ended treatment after six months, two people after nine months and one person stayed for a period of ten months. Table 1 gives a summary of the research sample's treatment before and after inpatient treatment (based on patients' self-report). Appendix 1 offers a more detailed idiosyncratic description of the participants (in line with Thurin & Thurin, 2007), comprising a brief clinical description, diagnoses and the change in outcome scores per patient[14].

Table 1. Treatment history of research sample

Treatment history	Before inpatient treatment	After inpatient treatment	At five-year follow-up
Participants	**n**	**n**	**n**
Outpatient psychotherapy	17	13	8
Psychiatrist	18	9	10
Medication	16	6	7
Psychiatric hospitalization	7	4	

Note. n = 22; various treatments per patient. Outpatient psychotherapy: often long-term, many different therapists and forms of talking-therapy or alternative treatments were consulted before inpatient treatment; creative therapy was increasingly consulted after inpatient treatment. Medication (in order of occurrence): antidepressants, sleep medication, anxiolytic, antipsychotic. Psychiatrist: if not specified, psychiatrist provided counselling.

[14] https://doi.org/10.5334/pb.432.s1

Instruments

Before the follow-up interview, participants completed a form asking for information about gender, age, occupation, medication use and psychotherapeutic treatment (before and after treatment).

The Outcome Questionnaire (OQ-45; Lambert et al., 1996; de Jong et al., 2007). The OQ-45 is one of the most frequently used self-report instruments for measuring patient clinical distress and well-being. The measure consists of three dimensions: subjective (symptomatic) distress (intra-psychic functioning), interpersonal relations and social role. Forty-five items are filled out on a 5-point Likert scale (0 = not at all; 4 = always). Sub-scores are calculated for the three sub-categories, the sum of which gives a total-score, ranging from 0 to 180; higher scores reflect greater distress. The Dutch version of the OQ-45 shows good internal consistency ($\alpha = 0.96$; total score), good test-retest reliability after two to three weeks ($r = 0.82$; total score for patient population) and sufficient concurrent validity (de Jong et al., 2007).

Semi-structured interview. The first author administered an adjusted version of the semi-structured **Client Change Interview** (CCI; Elliott, Slatick & Urman, 2001). The interview guide was constructed to evoke participants' experiences of the period before the inpatient treatment and the reasons for consulting the therapy centre, the changes they believe occurred during and after this treatment period, and what they believe influenced these changes. Interviews were audiotaped and transcripts were analysed.

Procedure

An explanatory sequential mixed-method study (Hesse-Biber, 2010) was conducted. The quantitative outcome scores and the narratives of former inpatients were analysed separately and integrated at the phase of interpretation, allowing for a complementation and deeper understanding of both the quantitative and qualitative outcome findings. This design can be

summarised as "quan → QUAL": a qualitative strand was added secondary to an existing quantitative study yet becoming the most important focus of the explanatory study (cf. Hesse-Biber, 2010).

The majority of the interviews were conducted at people's homes, a few were held in the therapy centre and one was administered at the faculty of Psychology and Educational sciences (Ghent University). Interviews lasted 100 minutes on average. An initial in-depth analysis was conducted on a sub-selection of 13 rich cases (displaying diversity in experiences), followed by an analysis of the remaining interviews (n = 11) using the initial themes as a compass. Minor adjustments were needed in order to encompass all narratives in a single final model of long-term outcome

Long-term quantitative outcome classification

Participants were classified in terms of reliable change and clinical significance based on four categories: clinically significant change (CS), reliable change (RC), no RC and deterioration (Jacobson & Truax, 1991). The outcome scores of the inpatient population were compared to Dutch norms (de Jong, Nugter, Lambert & Burlingame, 2008). In order to reach reliable change for the OQ-45 total score, a person must show a decrease in scores equal to or larger than 14. The cut-off between the clinical and nonclinical population for the Dutch OQ-45 is set at 55 (based on the test-retest reliability of 0.96; de Jong et al., 2007). We present all available quantitative outcome data (i.e., starting from the entire sample participating in the initial study) measured at the start of treatment (n = 43), at treatment ending (n = 27), at one-year (n = 29) and five-year follow-up (n = 22). To control for potential effects of drop-out, we compared the averages score on the OQ-45 (at the start of treatment) in the drop-out group (n = 9) and the remaining group (n = 34) by means of an independent samples t-test. Dependent samples t-tests were used to verify whether the outcome scores for the remaining group differed statistically

significant at the end of treatment, one-year follow up and five-year follow up compared to the start of treatment.

Qualitative analysis

A data-driven Thematic Analysis (Boyatzis, 1998; Braun & Clarke, 2006) based on principles of Grounded Theory (Glaser & Strauss, 1967), was conducted. This form of inquiry enabled the exploration of the phenomenon of long-term outcome in the participants' terminology and to identify themes in the data in a bottom-up manner. In our understanding of the first-person perspective, we adhere to a contextual idea: people give meaning to their experiences yet are influenced by the broader social context (Boyatzis, 1998; Braun & Clarke, 2006). The analysis focused on the macro-process of psychotherapy, i.e., a wide angle in contrast to the analyses of micro-processes (e.g., specific therapy effects). Moreover, the study aimed at investigating the subjective experience of several different participants (i.e., between-case variation) (Denzin & Lincoln, 2005).

Phases of analysis

Both authors coded the interviews separately; after each phase of analysis discussions were held until agreement was reached through a consensual process (cf. Hill, Thompson & Williams, 1997). The process of inquiry contained the following steps that were moved along in an iterative manner:

1. *Getting familiar with the narratives: emerging research question*

While conducting and reading the interviews, important individual differences among former inpatients' experiences became apparent. Subsequently, this variety among participants was incorporated in the analysis of experienced changes and RQ2 was further specified: *how can the explanatory factors help to understand individual differences between former patients' experienced changes?*

2. *Identifying relevant meaning units (initial sample, n = 13)*

Meaning units were identified (cf. Giorgi & Giorgi, 2003) by selecting chunks of text entailing all forms of experienced change (RQ1) and all mentioned influential factors in or outside treatment (RQ2). In this phase, we attempted to keep our perspective as broad as possible to prevent excluding relevant information too soon. Non-relevant parts, i.e., not dealing with subjects like well-being, experienced change or the therapy centre, were omitted.

3. *Providing codes for the selected meaning units (initial sample, n = 13)*

We attached keywords to text segments (Kvale & Brinkmann, 2009), systematically narrowing down the number of codes by integrating and renaming codes to the final version. For example, "being less insecure" and "more confidence" were merged together in a following step, resulting in the final category of "empowerment" (Mortelmans, 2011). RQ1 and RQ2 were coded for simultaneously, yet by using different indications or colours they were visually and thematically separated.

4. *Clustering codes to build core and subcategories (initial sample, n =13)*

The clustering of themes and factors was based on a thematic understanding of the perspective of all 13 participants and was conducted in several phases, in order to find the best-suited representations (e.g., the term "mollification" was chosen over "softening" or "appeasing" based on its definition "to soothe in temper or disposition"; Merriam-Webster online dictionary). At this phase, the first author reread the 13 interviews in order to verify whether the themes and factors still resonated with the participants' narratives.

5. *Preliminary taxonomy and explanatory model (initial sample, n = 13)*

Two separate preliminary models resulted: a taxonomy of experienced change (RQ1) and an explanatory model (RQ2). Both were developed to represent an overall representation as well as patients' individual differences. This resulted in twofold subcategories (e.g., "an altered self" consists of "empowerment" and "mollification").

6. *Coding the second part of the interviews using the initial categories as a compass*

Minor differences were found between the preliminary model and additional interviews and adjusted by adding subcategories (for example “A new perspective”) and renaming categories so terms would fit the uttered meanings of the entire sample.

7. *Resulting conceptual model of long-term outcome: merging models*

Ultimately, we agreed on merging both the taxonomy of change and explanatory factors into one overall model, as this allowed a clearer visual representation of the interrelations and influences as experienced by the participants. After this final phase, relevant citations were selected to provide rich examples.

Reliability and validity

In function of the reliability of the study, we tried to remain transparent about the entire process (Stiles, 1993) and acknowledge the influence of the perspective and background of the researchers. The first author conducted the interviews, making her personally involved with the participants and consequently also with the therapy centre. By avoiding preparatory knowledge (on the treatment setting and the quantitative outcome results), the influence of this personal involvement was countered as much as possible. The researchers’ personal interest in idiosyncratic experiences clearly navigated the focus of the study and analysis. By making this explicit, consequences of implicit guiding assumptions were controlled as much as possible (Creswell & Miller, 2000). In function of the validity of the study we worked in a systematic manner to form conclusions and interpretations (Stiles, 1993). Triangulation among researchers, several interviews (divided into two subsamples), and quantitative and qualitative indications of outcome were applied to gain different perspectives on the issue. The ultimate themes were formed by asking critical questions regarding codes and categories (Mortelmans, 2011). We attempted to stay open for any information coming from the narratives throughout the entire process (cf. non-framing interview format).

Ethics

Both the ethical committee of health institution Emmaus and the Ethical Committee of the University Hospital of Ghent University (Belgium) (registration number: B670201524637) approved the project. Every participant signed a written informed consent, for both the initial quantitative study and the follow-up inquiry.

Results

We first present the quantitative long-term outcome data from the start of therapy to five-year follow up and look into patients' perspectives on change and explanatory factors thereafter. Interpretations of the separate findings will be described in the result section; broader integrative conclusions and implications will be presented in the discussion.

Long-term quantitative outcome

Table 2 summarizes the change in average total scores on the OQ-45 over time, the standard deviations (SD) and range in scores (i.e., minimum and maximum score) as well as their meaning in comparison to a clinical and normal population. Table 3 shows the number of patients in the clinical and functional range, as well as CS change, RC, no change and deterioration.

The average total outcome score decreased from start to end of treatment, and slightly increased again during the follow-up period. The decrease in scores showed to be statistically significant at all time points (for $n = 34$) with an average decrease of 30 points (95% CI: 21.01 – 39.1; $p < 0.01$) from start to end of treatment, 23.2 points (95% CI: 14.38 – 32.1; $p < 0.01$) from start to one-year follow up and 18.7 points (95% CI: 7.68 – 29.69; $p < 0.01$) from start to five-year follow up. The average distress level in the drop-out group ($M = 100.56$; $SD = 12.053$) and remaining group ($M = 96.68$; $SD = 17.728$) did not differ significantly ($M = -3.879$; $SD = 6.287$; 95% CI: -16.577 – 8.818, $p = 0.274$); the decrease in scores thus does not seem influenced by the participants who dropped out from the study. Despite the significant diminution in severity of distress over time, the average score remains high in comparison to a

normal population. In terms of clinical significance, the changes in scores indicate the majority of participants have improved reliably, and this until five-year follow-up, yet only a minority reaches and remains the level of clinically significant change. The large range in scores further suggests a strong variety in the level of functioning among participants at all measurement points.

Table 2. The average, variation and meaning of outcome scores measured with the OQ-45 (total scores)

Average total score (OQ-45)	**Research sample**		**Clinical population**		**Normal population**	
	M	**SD (range)**	**Meaning**	**Percentile**	**Meaning**	**Percentile**
Start therapy	97	16.5 (60 – 144)	High	80-95	Very high	95-100
End therapy	67	22.2 (36 – 115)	Below average	20-40	High	80-95
One-year FU	73	22.7 (24 – 124)	Average (men) Below average (women)	40-60 20-40	Very high	95-100
Five-year FU	76.7	19.6 (30 – 116)	Average	40-60	Very high	95-100

Note. Research sample: n varies across measuring points. Clinical population: 628 men and 896 women; Normal population: 296 men and 511 women (de Jong et al., 2008).

Table 3. Clinical significance of outcome scores measured with the OQ-45 (total scores)

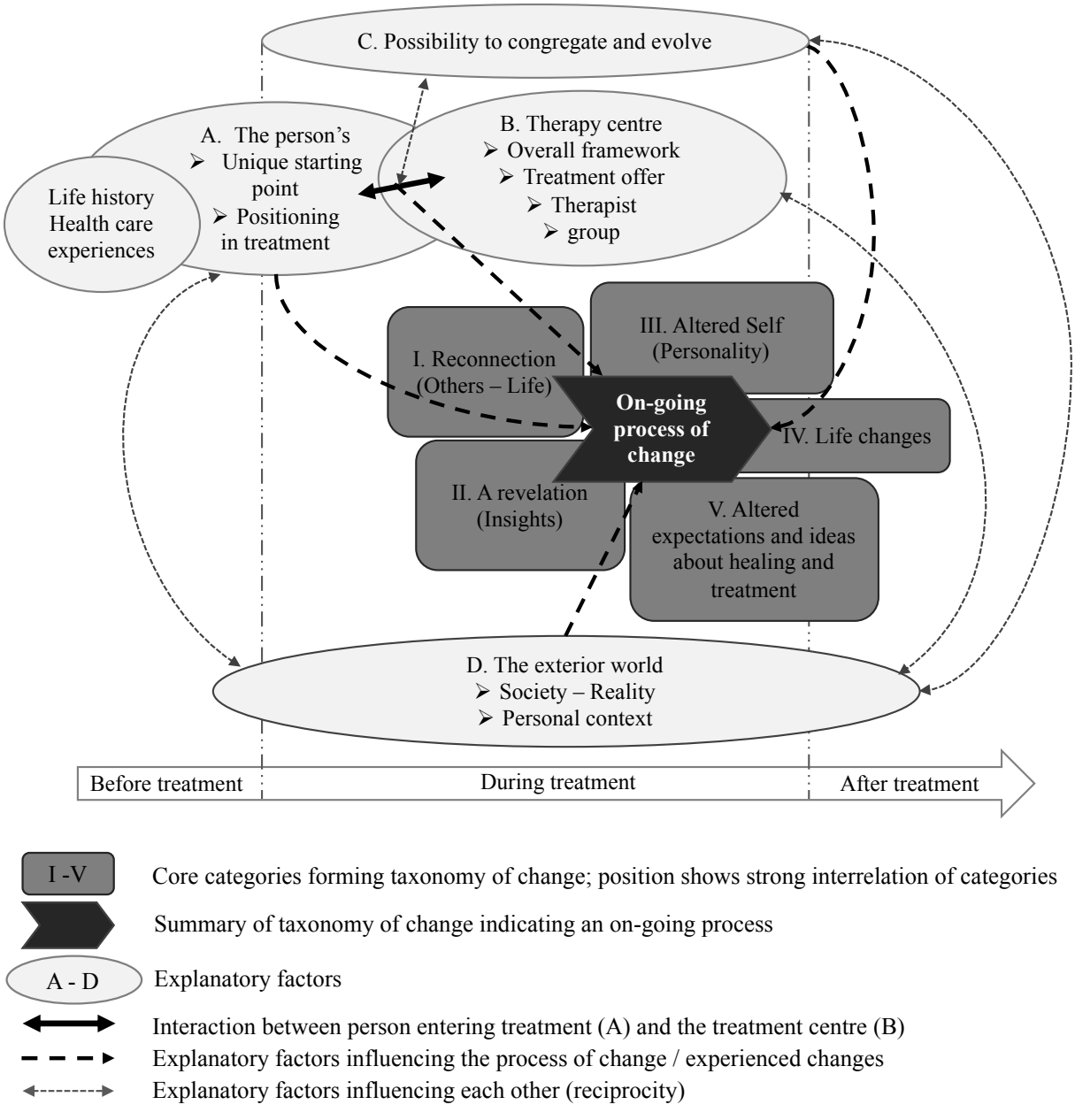
OQ-45 total score	**Start therapy**	**End therapy**	**One-year FU**	**Five-year FU**
	n	**n**	**n**	**n**
Participants	43	27	29	22
Functional	0	9	6	3
Clinical range	43	18	23	19
Change from start treatment				
CS		9	6	3
RC		12	15	11
No RC		5	5	6
Deterioration		1	3	2

Note. n varies due to the varying response rate at each measurement. Functional distribution: below clinical cut-off; clinical range: above the clinical cut-off; CS: reliable change and crossing the clinical cut-off; RC: reliable change only; No RC: criteria of statistical reliable change not met; Deterioration: reliable deterioration (reliable change in negative direction). The cut-off between clinical and non-clinical population for the OQ-45 = 55; reliable change on the OQ-45: difference in scores ≥ 14.

Conceptual model of long-term outcome of inpatient psychotherapy

Figure 1 presents the model of long-term outcome that comprises both a taxonomy of change and explanatory factors. In order to indicate variation in the prevalence of the categories and factors we rely on different phrasings: "(nearly) all" participants, indicates 90-100%, "most of" or "the majority" stands for 80-90%, "many" or "often" indicates 60-80%, "some(times)" or "others" means about 30-50%, and "a few" ranges around 10-20% of the participants.

Figure 1. Conceptual model of long-term outcome comprising the taxonomy of experienced changes (I – V) and explanatory factors (A – D)



Taxonomy of change: an on-going and non-linear process of existential changes

Table 4 presents the taxonomy of the experienced changes during and after inpatient treatment, differentiated in five categories that represent different levels of existential change: I. Reconnection to others and (the meaning of) life, II. A revelation, III. An altered self, IV. Life changes, and V. Altered expectations and ideas about recovery and treatment. In the conceptual model of long-term outcome presented in Figure 1, these changes are depicted along a timeline, indicating changes occurred at different moments in time. Without making causality claims, former patients described how the initial changes (i.e., I & II) have helped them engage in treatment and facilitated later changes (i.e., III, IV & V). In the resulting model, the experienced changes are positioned closely to each other, as they must be understood as strongly interrelated, i.e. changes coincided and influenced each another.

Table 4. Taxonomy of the experienced changes

Core and subcategories
I. Reconnection to others and (the meaning of) life
i. A feeling of belonging
ii. A new perspective
II. A revelation
i. Insight in self and difficulties
ii. Alternative ways of expressing emotions and thoughts
III. An altered self
i. A mollified self
ii. An empowered self
IV. Life changes
i. Concrete changes in dealing with or handling things in life
ii. Life-altering changes
V. Altered expectations and ideas about recovery and treatment
i. Recovery as on-going and fluctuating process
ii. Resignation versus disappointment

I. Reconnection to others and (the meaning of) life

Foremost, nearly all former patients experienced ***a feeling of belonging*** at the inpatient treatment centre. They described how they could recognise themselves and their difficulties in fellow patients, were accepted for who they are and felt worthy. Furthermore, patients gained a

new perspective on life, were filled with renewed energy and enthusiasm for life, and gained hope and courage to go on and overcome difficulties. Together these changes were interpreted as an experience of ***reconnection***, which is in stark contrast to the *disconnection*, i.e., exclusion, loneliness, failure and hopelessness, people experienced in society before entering inpatient treatment. For many, this feeling of reconnection was not only an important condition for engaging in therapy, but also meant a lasting change as such.

> 'What I remember very well is the aspect of being heard, being respected and treated as normal. I always had the feeling there was something wrong with me, that I was not okay, or it was not good enough, I had to try harder or they did not want me; I did not belong. At [the inpatient therapy centre] that all vanished. Now I feel at home anywhere I go.'

II. A revelation

Secondly, for nearly all participants, the treatment period lead to ***a revelation***. First of all, people mentioned "eye-openers" or ***insights in themselves and their difficulties*** as one of the most important changes.

> 'The biggest change for me was that, you think you know what your problem is, but you notice it can be something completely different. (…) At a certain point, you notice it goes much deeper.'

Also, people described how the various creative forms of therapy offered them ***alternative ways***, or even a first opportunity, to express emotions and thoughts; many continued to use these new ways after the treatment period (in the form of therapy, taking specific courses or hobbies).

> 'To the outside world I have always been the "ever-smiling person", I did not know anger and swallowed my frustration. Here [in the inpatient therapy centre] it changed a lot. I have learned to give words to my feelings, by first painting them on a canvas or by making something in clay.'

III. An altered self

At a later stage in treatment and to a lesser extent after treatment, patients experienced changes in their personality. On the one hand, many participants had become milder for themselves and others, showing a process of ***mollification***. Patients stated they have learned to acknowledge their own boundaries and accept difficulties that happened in the past, by taking them seriously.

> 'I have become softer, or that's what I am being told now by others. At my job, I now indicate my limits beforehand, as if I *allow* myself a private life now. (…) And I have learned it's okay to do 'nothing' for a while.'

On the other hand, many of the participants indicated they have become more independent, stronger and gained self-esteem, reflecting a feeling of ***empowerment***.

> 'I have learned to stand up for myself, to choose for myself. I have been able to take more distance from my parents, with the feeling: "it's okay; it's okay to think about yourself, to take yourself seriously".'

IV. Life changes

Experienced changes at the level of personality coincided with important ***concrete changes in dealing with or handling things in life***, increasingly taking place during and after treatment.[15]

> 'Since then [the inpatient treatment] I have always gone home [from work] on time and never took my computer with me. At five o'clock I could say: "it's time to go home, I'm out." Before, coming home meant turning on the computer and I could keep on working.'

Apart from changes in employment, also housing was often reconsidered and changes in family constitution and intimate relations (e.g., new partner or divorce) occurred. For some of the participants these changes were ***life-altering***, as the period of inpatient psychotherapy truly meant a turning point in life.

> 'My entire world has changed by coming here [the inpatient therapy centre]. I opened up here, my social life started and subsequently I met my husband. After years of invalidity and depending on my parents, my husband and I now

[15] Note that the difference between "an altered self" and "life changes" is specifically situated on the level of the experience and feeling of "being different" (e.g., different mind-set, attitude and character) versus the concrete changes or actions a person undertook resulting from or coinciding with changes in personality (cf. categories are strongly interrelated).

> live together and have a little toddler running around. I still feel tired, but I have people that rely on me now.'

V. Altered expectations and ideas about recovery and treatment

Nearly all of the participants disclosed they have been dealing with severe difficulties throughout their lives and none of them stated they were "cured". However, during and after the treatment period a ***change in ideas and expectations about recovery and treatment*** occurred. The initial idea that after a couple of months all problems would be solved, changed for most patients to perceiving recovery as ***an on-going and fluctuating process***. Apart from the many valuable changes, and independent from the situation at the time of the interview, many participants also experienced a breakdown or recurrent difficulties.

For most of the participants, seeing recovery as an on-going and rather fluctuating process functioned as a ***resignation***, leading to adjusted (i.e., more realistic) expectations and understanding of their own process of recovery, subsequently influencing concrete changes in life and coinciding with changes in personality (III & IV). A few participants, however, seemed to hold on to a more fixed idea about recovery or had an ideal image of treatment, eventually resulting in a ***disappointment***, in themselves or in psychotherapy, as these (sometimes high) expectations were not met.

> 'I believed all my problems would be solved: I would easily find work and build new social contacts, I would have a better relationship with my parents. But I expected too much. My image about the world and future had become too idealistic.'

Explanatory factors of change: understanding individual differences

Figure 1 depicts four broad factors in patients' narratives that have explanatory value with respect to the experienced changes and their interrelation: A. The person's unique starting point and positioning in treatment, B. Central elements of the therapy centre (in relation to the person), C. The possibility to congregate and evolve, and D. The relation to the exterior world.

These factors must be understood as evolving throughout and after the treatment period and are thus also depicted along the timeline in Figure 1.

A. The person's unique starting point and positioning in treatment

Important differences were observed among former patients regarding *when or where* in the course of their lives and health care trajectory they had entered inpatient treatment, making up ***the person's unique starting point***. Whether the inpatient treatment followed previous psychotherapeutic experiences or not, consequently influenced the process of inpatient therapy, the extent to which insights were gained and changes in personality and life could occur.

> 'At that moment [in inpatient treatment] I was really ready for it, and I really wanted to do something about it. A bit opposed to the years before, I think. When I was hospitalized in psychiatry they mainly had to keep me alive. This [the inpatient therapy centre], however, has been my final hospitalization.'

Secondly, former patients described different ways of ***positioning oneself in treatment*** that could be summarized as: dependent-passive, assertive-active, overactive, and independent – keeping a distance. These four positions are neither mutually exclusive nor fixed yet present a predominant way of engaging in treatment; Table 5 provides examples of every position. Whether someone displayed a high need of guidance, engaged actively in treatment, was preoccupied with gaining instant effects, or had difficulties opening up in treatment (e.g., due to trust issues), influenced the treatment process differently and subsequently also the experienced changes.

Table 5. Examples of different positions in treatment

Position in treatment:	Example excerpt
***Dependent-passive**: high need of guidance, support and initiative from therapists*	'It was hard when the initiative had to come from me. Maybe it's my personality; I often need a lead, as I often tend to lean on others. And maybe it was the plan of the therapists that I would take initiative and stand up for myself. But having to create things myself in therapy was hard.'
***Assertive-active:** highly engaged in treatment process, taking own initiative*	'Being here [in the inpatient therapy centre] was a unique opportunity I had to grasp with my both hands. I chose for myself, I wanted to become a happier person, so I engaged in everything they offered me. At a certain point, I even asked for an extra form of therapy.'
***Over-active:** preoccupied with "working hard" and time-efficiency of treatment*	'I had prepared myself for [the inpatient] treatment. I had made drawings with all my characteristics on it, I wanted to know all of it, I did not want to do nothing there [in the inpatient therapy centre]; I wanted to work on myself.' 'When I was drawing in therapy I thought: "I have to be at my department, I have to lead my team, my colleagues!"'
***Independent-keeping a distance:** avoiding participation or inhibited to fully engage*	'I've realized, I'm not very good at therapy... Personal conversations with people I don't know so well are hard, it takes a while before I trust a person. I'm very secretive. I constantly censor myself and consider 'what can I share and what not' and "shouldn't I be solving this on my own".'

B. Central elements of the therapy centre (in relation to the person)

A reciprocal arrow between the person and therapy centre in Figure 1 signifies the importance of the interaction between both factors. Specifically regarding the interaction of the patient and the ***overall approach*** of the therapy centre, people who described a more *dependent-passive* and *independent-keeping a distance* position stated to have encountered the most difficulties in displaying autonomy and self-reliance. Patients with an *assertive-active* position seemed to engage more easily in a process of self-exploration, while people in an *over-active* position encountered difficulties with the dimension of taking time for oneself (e.g., due to a feeling of not working hard enough); see Table 5.

Accordingly, the extent to which the treatment was experienced as well adjusted to personal needs differed from participant to participant, yet all mentioned at least one therapy in their ***treatment package*** they had found helpful (cf. adjusted therapy programme). The connection with the ***therapist*** (and by extension the personal mentor and psychiatrist) and the ***group*** appeared crucial for changes to occur.

> ‘It was hard for me to open up, and most of them [the therapists] let me be, but there was one therapist who saw through me. He said: “you are here long enough, you can make it a little bit harder on yourself”, and I thought “he is right”. The fact that he did not tolerate my attempts to hide [was important]’

C. The possibility to congregate and evolve

The person, the treatment and their interaction must not be understood as fixed. The possibility to evolve towards a coming together of both (i.e., harmony and accordance) appeared essential to understand why changes came about or not. For instance, many described how they have evolved in the position of being in treatment (e.g., from passive to active; overcoming uncertainty and distrust), or how therapy was altered to reach a patient’s (changing) needs (i.e., via therapists’ interventions or by changing a therapy programme). A rift in the relation or evolution, on the other hand, could lead to the undermining of further treatment and change.

> ‘After nine months, they [the staff] told me: “we cannot continue to work with you, we are stuck.” And I knew it myself. They wanted me to participate actively but my reaction was to retreat, to build a wall around me. And they were right, no one could get in.’

D. The relation to the exterior world

Participants described their experiences explicitly in comparison and in relation to ***society*** or ***reality*** outside the therapy centre. To many, the period in the inpatient treatment centre offered a healing experience and buffer against hostile encounters (e.g., discrimination, stigmatization, high demands and pressure), although some also suffered from this contrast with the outside world. [16]

> ‘I miss the warmth of the therapy centre; people listened to you, it was okay to say “I’m not feeling very well today”. In society, this is just not tolerated, if you state you’re not doing okay you are pushed aside; you always need to be strong.’

[16] The term “reality” was taken from patients’ own description. A distinction between the “outside world” (broadly termed as “the society”) and the “inside world” of the treatment centre was strongly made. This reality or outside world was mostly seen as “cold” and in stark contrast with the safe and caring environment of the treatment centre.

Furthermore, the ***personal context*** (i.e., family, social and professional context) had an important influence on the treatment and change process. For some, this context was a reason for being hospitalized and distance as such was salutary; to others, the context was left aside too much during the treatment period. The extent to which, for instance, family, social life and further career planning had been integrated into the treatment period seemed an important facilitator of change and transition to daily life after ending treatment. For some patients, this was however beyond their possibilities at this point in their treatment trajectory (cf. unique starting point; A). Further, family at home sometimes meant an important reason to prematurely end therapy.

> 'My daughter asked me to come home again, as I had promised my treatment would maximally take one year. But actually, my treatment wasn't entirely finished then. Now I regret that [premature treatment termination].'

Discussion

Long-term outcome findings show improved well-being for the majority of former patients and this until five to six years after ending inpatient treatment. Yet, in order to obtain a deeper understanding and contextualization of these findings, long-term outcome scores were complemented by the patients' own perspective.

The conceptual model based on the qualitative analysis of patients' narratives indicates long-term outcome can be understood as an ongoing and fluctuating process of predominantly existential changes. These changes are furthermore strongly interrelated and were described as coinciding and influencing each other over time (cf. Binder et al., 2009; Ekroll & Rønnestad, 2016). In line with previous (mostly outpatient) qualitative research, participants in this study mentioned changes in the domains of the self (in terms of personality), interpersonal relations (here more broadly described as the exterior world), and specific life conditions (Klein & Elliott, 2006; Midgley et al., 2006; Nilsson et al., 2007). Concrete life changes seem to correspond to behavioural changes mentioned in other studies, while alternative ways of

expressing emotions and thoughts may be seen as different ways of coping (Ekroll & Rønnestad, 2016; Midgley et al., 2006; Nilsson et al., 2007). Accordingly, we observed a decrease in treatment intensity (i.e., medication, counselling, and hospitalization) following the inpatient treatment, while creative forms of therapy and other activities increased. Central to our findings is that none of the participants considered themselves as "cured" during or after treatment, yet nearly all emphasized the importance of the inpatient treatment for obtaining fundamental changes. This seems to correspond to the observation that only a minority of patients reached clinically significant change in outcome scores, although the majority showed reliable improvement compared to their initial level of functioning. Interestingly, participants in this study evaluated their process of recovery in light of their own personal life story while rejecting the normative standards of society. The inpatient treatment centre offered a healthier and healing experience that was strongly contrasted with the outside world. Whilst previous qualitative studies found changes in terms of accepting own limits and weaknesses (Kühnlein, 2011; Nilsson et al., 2007), unique for this study seems to be patients' different perspective on recovery and treatment.

This finding touches upon the ethical question of how recovery, functioning and dysfunctioning are conceptualized. In accordance with the altered perspective of the participants in this study, a broader view on recovery can be noticed in the organisation of (mental) health care: a dichotomous understanding of "health" versus "illness" is increasingly replaced by a definition of "positive health" (Delespaul et al., 2016; Huber et al., 2011; Slade & Longden, 2015). The latter emphasizes the possibilities and abilities, *given* psychological, physical and social challenges and is considered more in line with clinical practice (Delespaul et al., 2016). For most of the participants, this different appraisal of recovery offered resignation, while fixated ideals were met with disappointment. Conceptualizations of dysfunctioning and recovery also provide the (often implicit) assumptions underlying the

measurements in psychotherapy and outcome research (Imel & Wampold, 2008; Rieken & Gelo, 2015; Slife, 2004). From a positive perspective on recovery, the individual process and evaluation becomes central and comparison to others is only possible to a certain extent (Slade & Longden, 2015). The research field faces the challenge of adapting to the shifting perspective on outcome and the growing body of idiosyncratic research (quantitative and qualitative) is clearly promising in that respect. The current study aimed at contributing to these developments in the context of inpatient psychotherapy.

Implications for further research on outcome

Our findings highlight the multiplicity of long-term outcome, substantiating the relevance of a broader conception and measurement of treatment effects. The complex interplay of the patient, treatment, external context and evolution over time further indicates that experienced changes cannot be explained by single or fixed factors. In this respect, our results support a dynamic understanding of the patient-therapist relationship as suggested by Clarkin and Levy (2004), although "the therapist" requires a broader interpretation in our study (i.e., elements of the treatment centre). Despite its residential nature, the influence of the context outside therapy must also be considered in this dynamic patient-treatment relationship (cf. McLeod, 2016). Regarding the patient, our study points at the importance of a patient's starting point in treatment (influenced by life experiences and previous health care), which relates to the concept of patient characteristics that is broadly considered the single most important factor for explaining treatment effects (for an overview, see Bohart & Wade, 2013). While the patients' perspective is often explored after treatment, it is arguable that a more in-depth understanding of the patient's starting position can enhance knowledge on tailoring psychotherapy to individual needs (Norcross & Wampold, 2011). Qualitative analysis of patient's experience before treatment could mean a valuable contribution here. Further, whereas most research assesses patients' expectations statically, our study shows the potential relevance of evaluating

treatment expectations as changing over time (Greenberg, Constantino & Bruce, 2006). Longitudinal prospective research would enable to look at individual trajectories of change in inpatient treatment (Bohart & Wade, 2013). The different positions in treatment that were found in this study (i.e., dependent, assertive, over-active and avoidant), seem to relate to several characteristics described in psychotherapy literature, like a patient's "state of readiness", level of engagement and possibility to establish good therapeutic alliance (Bohart & Wade, 2013). Further research could focus on how patients' positioning in treatment evolves and can be facilitated during inpatient care (cf. individualized treatment programmes). The extent to which these different positions can be observed elsewhere requires replication in other settings. Likewise, our findings suggest a feeling of reconnection might be essential for patients to engage in treatment (cf. Ekroll & Rønnestad, 2016). The question can be raised whether this feeling is unique for this setting, specific for inpatient long-term treatment, or can also be observed in outpatient care and across different groups of patients.

Strengths, limitations and future directions

The longitudinal design of this study enabled to outline the sustainability of changes over time; the idiosyncratic perspective of former inpatients allowed understanding these changes beyond symptomatic evolutions. Both can be considered important strengths of this study. High levels of complexity and diversity characterize the field of inpatient psychotherapy, and likewise, the treatment centre in this study cannot be considered identical to other inpatient settings (cf. Kösters et al., 2006). Nonetheless, several elements of residential care (e.g., living together on a ward; different forms of group and creative therapy) can be considered shared among inpatient facilities and therefore support transferability to a certain extent (e.g., Leszcz et al., 1985). Still, further research on inpatient settings is warranted to deepen the understanding of this particular therapy context. In the present study, we aimed to contribute to this inquiry by incorporating all available outcome scores and all 22 interviews, resulting in an overall-experience and

general model of inpatient treatment. Investigating patients' experiences of more specific aspects of inpatient treatment (e.g., the role of creative or group therapy), during or shortly after ending treatment, could complement our findings.

The participants in this study resemble a homogenous and local (Caucasian, Flemish) group of former patients. Studies that focus on specific (e.g., cultural, ethnic) or more diverse groups of patients (e.g., gender) could therefore offer an important contribution. Our participants also formed a sub-group of patients that were willing to participate in the study. This could imply that more dissatisfied people were not heard and that therefore relevant knowledge remains unexposed. Although we have incorporated negative experiences in our analyses (cf. Midgley et al., 2006; Nilsson et al., 2007), qualitative research focusing on negative outcomes remains scarce whereas it could offer an important contribution to the understanding of (inpatient) psychotherapy (e.g., Werbart, von Below, Brun & Gunnarsdottir, 2015). In line with this, effects of social desirable answers cannot be fully excluded (Thurin & Thurin, 2007).

Many external influences could have shaped former patients' perspectives in the period of five to six years after ending treatment. Due to its naturalistic character, our study is limited in terms of standardization and control. For instance, the research sample is characterized by diverse clinical problems (e.g., diagnoses) unlike what is typical in standard outcome studies (Westen et al., 2004). The model of long-term outcome must thus be considered preliminary and we wish not to make strong causal claims regarding the effectiveness of treatment. Rather, the strength of a naturalistic research design is the clinically representative study of a particular context and population of patients, resulting into clinically valuable knowledge (Leichsenring, 2004). Finally, following Strupp and Hadley (1977), we agree that the chosen first-person perspective is only one perspective on outcome (e.g., compared to the therapist or societal perspective), and likewise only sheds light on certain aspects, while leaving others untouched.

Nevertheless, studying the patients' perspective offers major contributions to the clinical meaningfulness of research findings and understanding of psychotherapy outcome.

Conclusion

This study offers support to the idea that, instead of relying on a unified and delineated definition and measurement of outcome or recovery, research and practice can greatly benefit from acknowledging the inherent complexity, multiplicity and diversity of psychotherapy outcome that require the incorporation of multiple methods and perspectives.

References

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders. Text Revision (4th ed.).* Washington, DC: American Psychiatric Association

APA Presidential Task Force on Evidence-Based Practice (2006). Evidence-based practice in psychology. *American Psychologist, 61*(4), 271-285. Retrieved from http://www.ovid.com/site/index.jsp

Blanton, H., & Jaccard, J. (2006). Arbitrary metrics in psychology. *The American Psychologist*, *61*(1), 27–41. doi: 10.1037/0003-066X.61.1.62

Binder, P. E., Holgersen, H., & Nielsen, G. H. (2009). What is a "good outcome" in psychotherapy? A qualitative exploration of former patients' point of view. *Psychotherapy Research (iFirst)*, 1-10. doi: 10.1080/10503300903376338

Binder, P. E., Schanche, E., Holgersen, H., Nielsen, G. H., Hjeltnes, A., Stige, S. H., … C., Multu. (2016). Why do we need qualitative research on psychological treatments? The case for discovery, reflexivity, critique, receptivity, and evocation. *Scandinavian Psychologist*, *3*, 1–22. doi: 10.15714/scandpsychol.3.e8

Bohart, C. A., & Wade, A. M. (2013). The client in psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behaviour change* (6th Ed.) (pp. 219-257). Hoboken, New Jersey: John Wiley & Sons, Inc.

Boyatzis, R. E. (1998). *Transforming qualitative information: thematic analysis and code development.* London, England: Sage

Braakmann, D. (2015). Historical paths in psychotherapy research. In O. C. G. Gelo, A. Pritz & B. Rieken (Eds.), *Psychotherapy research: foundations, process and outcome* (pp. 39-66). Vienna, Austria: Springer.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3,* 77-101. doi: 10.1191/1478088706qp063oa

Castonguay, L., Barkham, M., Lutz, W., & McAleavey, A. (2013). Practice-oriented research: Approaches and applications. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behaviour change* (6th Ed.) (pp 85-133). Hoboken, New Jersey: John Wiley & Sons, Inc.

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 behaviour change* (5th Ed.) (pp. 194-226). Hoboken, New Jersey: John Wiley & Sons, Inc.

Creswell, J. W., & Miller, D. L. (2000). Determining validity in qualitative inquiry. *Theory into practice*, *39*(3), 124-130. Retrieved from https://people.ucsc.edu

de Jong, K. & Nugter, M. A., Polak, M.G., Wagenborg, J.A.E., Spinhoven, P., & Heiser, W.J. (2007). The Outcome Questionnaire (OQ-45) in a Dutch population: a cross-cultural validation. *Clinical Psychology & Psychotherapy, 14,* 288-301. doi: 10.1002/cpp.529

de Jong, K., Nugter, M. A., Lambert, M. J., & Burlingame, G. M. (2008). Handleiding voor afname en scoring van de Outcome Questionnaire OQ-45.2. Salt Lake City

Delespaul, P., Milo, M., Schalken, F., Boevink, W., & van Os, J. (2016). *Goede GGZ! Nieuwe concepten, aangepaste taal en betere organisatie.* Netherlands: Diagnosis publishers

De Los Reyes, A., Kundey, S. M. A., Wang, M. (2011). The end of the primary outcome measure: A research agenda for constructing its replacement. *Clinical Psychology Review, 31*, 829-838. doi: 10.1016/j.cpr.2011.03.011

Denzin, N. K., & Lincoln, Y. S. (2005). *The sage handbook of qualitative research.* Thousand Oaks, CA: Sage

Ekroll, V. B., & Rønnestad, M. H. (2016). Processes and changes experienced by clients during and after naturalistic good-outcome therapies conducted by experienced psychotherapists. *Psychotherapy Research, 3307,* 1-19.
doi.: 10.1080/10503307.2015.1119326

Elliott, R., Slatick, E., & Urman, M. (2001). Qualitative change process research on psychotherapy: Alternative strategies. In J. Frommer & D. L. Rennie (Eds.), *Qualitative psychotherapy research: Methods and methodology* (pp. 69–111). Lengerich, Germany: Pabst Science.

Englebert, J., Follet, V., Valentiny, C. (2017). Anorexia Nervosa and first-person perspective: altruism, family system and body experience. *Psychopathology, 51*, 24-30. DOI: 10.1159/0004852629

Englebert, J., Stanhellini, G., Valentiny, C., Fuchs, T., & Sass, L. (2018). Hyper-reflexivité et perspective en première personne: un apport décisif de la psychopathologie phénoménologique contemporaine à la comprehension de la schizophrénie. *L'évolution psychiatrique, 83*, 77-85. doi : 10.1016/j.evopsy.2017.07.004

Gaudiano, B. A., & Miller, I. W. (2013). The evidence-based practice of psychotherapy: Facing the challenges that lie ahead. *Clinical Psychology Review, 33*, 813–824. Retrieved from http://www.womha.com

Giorgi, A., & Giorgi, B. M. (2003). The descriptive phenomenological psychological method. In P. M. Camic, J. E. Rhodes & L. Yardley (Eds.), *Qualitative research in psychology: expending perspectives in methods and design* (pp. 243-274). Washington, DC: American Psychological Association

Glaser, B. G., & Strauss, A. L. (2009). *The discovery of grounded theory: Strategies for qualitative research*. Hoboken, New Jersey: Transaction publishers

Greenberg, R. P., Constantino, M. J., & Bruce, N. (2006). Are patient expectations still relevant for psychotherapy process and outcome? *Clinical psychology review*, *26*(6), 657-678. doi: 10.1016/j.cpr.2005.03.002

Hesse-Biber, S. N. (2010). *Mixed methods research, merging theory with practice.* Guilford press, New York

Hill, C. E., Chui, H., & Baumann, E. (2013). Revisiting and reenvisioning the outcome problem in psychotherapy: An argument to include individualized and qualitative measurement. *Psychotherapy*, *50*(1), 68–76. doi: 10.1037/a0030571

Hill, C. E., & Lambert, M. J. (2004). Methodological issues in studying psychotherapy processes and outcomes. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behaviour change* (5th Ed.) (pp. 84-135) Hoboken, New Jersey: John Wiley & Sons, Inc

Hill, C. E., Thompson, B. J., & Williams, E. N. (1997). A guide to conducting consensual qualitative research. *The counselling psychologist, 25*(4), 517-572. DOI: 10.1177/0011000097254001

Huber, M., Knottnernus, J. A., Green, L., van der Horst, V., Jadad, A. R., Kromhout, D., … Smid, H. (2011). How should we define health? *British Medical Journal,* 1–3. doi: 10.1136/bmj.d4163

Imel, Z., & Wampold, B. (2008). The importance of treatment and the science of common factors in psychotherapy. In S. D. Brown & R. W. Lent (Eds.), *Handbook of counselling psychology (4th Ed.).* (pp 249-262). Hoboken, New Jersey: John Wiley & Sons, Inc.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12-19. doi: 10.1037//0022-006X.59.1.12

Kazdin, A. E. (2006). Arbitrary metrics: implications for identifying evidence-based treatments. *The American Psychologist*, *61*(1), 42–71. doi: 10.1037/0003-066X.61.1.42

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

Kiesler, D. J. (1966). Some Myths of Psychotherapy Research and the Search for A Paradigm. *Psychotherapy Research*, *5*(2), 110-136. doi: 10.1080/10503309512331331186

Klein, M. J., & Elliott, R. (2006). Client accounts of personal change in process experiential psychotherapy: A methodologically pluralistic approach. *Psychotherapy Research, 16*, 91-105. doi: 10.1080/10503300500090993

Kösters, M., Burlingame, G. M., Nachtigall, C., & Strauss, B. (2006). A meta-analytic review of the effectiveness of inpatient group psychotherapy. *Group Dynamics: Theory, Research, and Practice*, *10*(2), 146–163. doi: 10.1037/1089-2699.10.2.146

Kühnlein, I. (1999). Psychotherapy as a Process of Transformation: Analysis of Post-therapeutic Autobiographic Narrations. *Psychotherapy Research*, *9*(3), 274–287. doi: 10.1080/10503309912331332761

Kvale, S., & Brinkmann, S. (2009). *Interviews: Learning the craft of qualitative research interviewing*. Los Angeles, CA: Sage

Lambert, M. J., Hansen, N. B., Umphress, V., Lunnen, K., Okiishi, J., Burlingame, G. M., & Reisinger, C. W. (1996). *Administration and scoring manual for the OQ-45.2*. East Setauket, NY: American Professional Credentialing Services.

Leichsenring, F. (2004). Randomized controlled versus naturalistic studies: A new research agenda. *Bulletin of the Menninger Clinic, 68*(2)*,* 137-151 doi:10.1521/bumc.68.2.137.35952

Levitt, H. M., Stanley, C. M., Frankel, Z., & Raina, K. (2005). An evaluation of outcome measures used in humanistic psychotherapy research: Using thermometers to weigh oranges. *The Humanistic Psychologist, 33*, 113–130. doi: 10.1207/s15473333thp3302_3

Leszcz, M., Yalom, I. D., & Norden, M. (1985). The value of inpatient group psychotherapy: Patients' perceptions. *International Journal of Group Psychotherapy*, *35*(3), 411-433. doi: 10.1080/00207284.1985.11491425

McLeod, J. (2011). *Qualitative research in counselling and psychotherapy*. London, England: Sage.

McLeod, J. (2016). *Using research in counselling and psychotherapy.* London, England: sage

Midgley, N., Target, M., & Smith, J. (2006). The outcome of child psychoanalysis from the patient's point of view: a qualitative analysis of a long-term follow-up study. *Psychology and Psychotherapy*, *79*, 257–269. doi: 10.1348/147608305X52694

Midgley, N., Ansaldo, F., & Target, M. (2014). The meaningful assessment of therapy outcomes: incorporating a qualitative study into a RCT evaluating the treatment of adolescent depression. *Psychotherapy 51*(1), 128-137. doi: 10.1037/a0034179

Mortelmans, D. (2011). *Kwalitatieve analyse met Nvivo.* Leuven, Belgium: Acco

Nilsson, T., Svensson, M., Sandell, R., & Clinton, D. (2007). Patients' experiences of change in cognitive-behavioural therapy and psychodynamic therapy: a qualitative comparative study. *Psychotherapy Research*, *20*, 37-41.
doi: 10.1080/10503300601139988

Norcross, J. C., & Wampold, B. E. (2011). What works for whom: Tailoring psychotherapy to the person. *Journal of Clinical Psychology*, *67*(2), 127-132. doi: 10.1002/jclp.20764

Rieken, B., & Gelo, O. C. G. (2015). The philosophy of psychotherapy science: mainstream and alternative views. In O. C. G. Gelo, A. Pritz, & B. Rieken, (Eds.), *Psychotherapy research: foundations, process and outcome* (pp. 67-92). Vienna, Austria: Springer

Slade, P. M., & E., Longden, (2015). Empirical evidence about recovery and mental health. *Bio Med Central Psychiatry*, 1-14. doi: 10.1186/s12888-015-0678-4

Slife, B. D. (2004). Theoretical challenges to therapy practice and research: the constraint of naturalism. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behaviour change* (5th Ed.) (pp. 44-83). Hoboken, New Jersey: John Wiley & Sons, Inc.

Stiles, W. B. (1993). Quality Control in Qualitative Research. *Clinical psychology Review, 13,* 593-618. doi: 10.1016/0272-7358(93)90048-Q

Stiles, W. B., Shapiro, D. A., & Harper, H. (1994). Finding the way from process to outcome: blind alleys and unmarked trails. In R. L. Russell (Ed.), *Reassessing psychotherapy research.* (pp. 65–113). New York, NY: Guilford Press

Stiles, W. B., & Shapiro, D. A. (1989). Abuse of the drug metaphor in psychotherapy process-outcome research. *Clinical Psychology Review*, *9*(4), 521-543. doi: 10.1016/0272-7358(89)90007-X

Strupp, H. H., & Hadley, S. W. (1977). A tripartite model of mental health and therapeutic outcome: With specific reference to negative effects in psychotherapy. *American Psychologist, 32,* 187–196. doi:10.1037/0003- 066X.32.3.187

Thurin, J.-M., & Thurin, M. (2007). *Évaluer les psychothérapies, méthodes et pratiques.* Dunod

Tschuschke, V., & Dies, R. R. (1994). Intensive analysis of therapeutic factors and outcome in long-term inpatient groups. *International Journal of Group Psychotherapy, 44*(2), 185-208 doi: 10.1080/00207284.1994.11490742

Valkonen, J., Hanninen, V., & Lindfors, O. (2011). Outcomes of psychotherapy from the perspective of the users. *Psychotherapy Research: Journal of the Society for Psychotherapy Research*, *21*(2), 227-240. doi: 10.1080/10503307.2010.548346

Westen, D., Novotny, C. M., & Thompson-Brenner, H. (2004). The Empirical Status of Empirically Supported Psychotherapies: Assumptions, Findings, and Reporting in Controlled Clinical Trials. *Psychological Bulletin, 130*(4), 631–663. doi: 10.1037/0033-2909.130.4.631

Werbart, A., von Below, C., Brun, J., & Gunnarsdottir, H. (2015). “Spinning one’s wheels”: Nonimproved patients view their psychotherapy. *Psychotherapy Research*, *25*(5), 546-64. Doi: 10.1080/10503307.2014.989291

III. What "Good Outcome" Means to Patients: Understanding Recovery and Improvement in Psychotherapy for Major Depression from a Mixed-Methods Perspective

Abstract

This study explored the meaning of "good outcome" within and beyond the much-used statistical indices of clinical significance in standard outcome research as developed by Jacobson and Truax (1991). Specifically, we examined the experiences of patients marked as "recovered" and "improved" following CBT and PDT for major depression. A mixed methods study was conducted using data gathered in an RCT, including patients' pre- post outcome scores on the BDI-II and post treatment Client Change Interviews. We selected 28 patients who showed recovery and 19 patients who showed improvement in self-reported depression symptoms. A grounded theory analysis was performed on patients' interviews, ultimately resulting in a conceptual model of "good outcome". From patients' perspectives, good outcome can be understood as feeling empowered, finding personal balance and encountering ongoing struggle, indicating an ongoing process and variation in experience. The JT classification of "good outcome" could not account for the (more pessimistic) nuances in outcome experiences, especially for "improved" patients, and did not grasp the multi- dimensional nature of outcome as experienced by patients. It is recommended that statistical indications of clinical meaningfulness are interpreted warily and ideally contextualized within personal narratives. Further research on the phenomenon of change and good outcome is required, aiming at integrating multiple perspectives and methods accordingly the multi-dimensional phenomenon under study[17].

[17] De Smet, M. M., Meganck, R., De Geest, R., Norman, U. A., Truijens, F., & Desmet, M. (2019). What "Good Outcome" Means to Patients: Understanding Recovery and Improvement in Psychotherapy for Major Depression From a Mixed-Methods Perspective. *Journal of Counseling Psychology.* http://dx.doi.org/10.1037/cou0000362

Introduction

The clinical meaningfulness of outcome findings is a central concern for outcome research and the field of psychotherapy and counseling at large. Randomized Controlled Trials (RCTs) especially, face the challenge of increasing their value for clinical practice, as their deficient ecological validity (i.e., whether findings apply outside of the research setting) is widely recognized (Carey & Stiles, 2015). Despite the abundance of evidence supporting the efficacy of psychotherapy, the clinical meaning of significant outcome remains unclear (Cuijpers et al., 2014; Kazdin, 1999; 2008).

In the history of attempts to increase the clinical relevance of findings from outcome research, the development of the statistical concept of clinical significance has been of major importance (Lambert & Ogles, 2009; Ogles, Lunnen & Bonnesteel, 2001). Clinical significance addressed the need for indicators of the clinical relevance of measured changes, which is not covered by typical statistical tests of significance and effect sizes (Cuijpers et al., 2014; Lambert, Hansen, & Bauer, 2008). Moreover, it shifted the focus from between-group mean-level evaluations of change, to clinical significance for individual patients (Jacobson, Follette, & Revenstorf, 1984; Kazdin, 1999; Ogles et al., 2001). Various methods of clinical significance have been developed for instance, "social validation" which relies on subjective evaluation (i.e., by significant others) or normative comparison (Kazdin, 1977). The most widely used method was designed by Jacobson and colleagues (Jacobson et al., 1984) - later refined by Jacobson and Truax (i.e., the "JT method", 1991; Ogles et al., 2001) - and is considered the classification of choice (Atkins, Bedics, McGlinchey & Beauchaine, 2005; Lambert et al., 2008).

The JT classification relies on a twofold evaluation: Reliable change in pre-to-post treatment scores and evolution from dysfunctional to functional distribution (i.e., passing the clinical cut-off). Accordingly, patients are classified into four possible outcome categories: "Recovery" (reliable change and passing the clinical cut-off), "Improvement" (reliable change but remaining in the clinical range), "Non-improvement" (no reliable change) or

"Deterioration" (reliable change in the negative direction). In psychotherapy outcome research, the category "recovery" is considered the gold standard for determining clinically significant change (Lambert et al., 2008; Ogles et al., 2001), although in practice the category "improvement" is viewed as an indicator of good outcome as well (e.g., Ekroll & Rønnestad, 2017; McElvaney & Timulak, 2013).

Even though the focus on clinical significance provided an important contribution to the interpretation of statistical outcome evaluation, it does not answer the question whether measured outcome represents changes that are actually meaningful to patients in therapy. A possible discrepancy between clinically significant changes based on pre-post, mainly symptom-based (Braakmann, 2015; Kazdin, 1999) outcome measures and patients' real-life experiences has been problematized (Blanton & Jaccards, 2006; Kazdin, 2001). In the past, this has led to the development and inclusion of quality of life measures (Gladis, Gosh, Dishuk, & Crits-Cristoph, 1999), indices of general well-being (e.g., Beckstead et al., 2003) and personalized questionnaires (Elliott, Mack, & Shapiro, 1999). Nevertheless, as RCTs typically rely on a *single* primary outcome measure to define treatment efficacy and consequently to calculate the clinical significance of patients' outcome (De Los Reyes, Kundey & Wang, 2011), further conceptual clarity on the meaning of the statistical concept of clinical significance in relation to the impact of therapy on patients' lives is needed (Atkins et al., 2005; Kazdin, 2001).

In this regard, qualitative inquiry of patients' perspectives is increasingly put forward as a contribution to outcome studies and RCTs in particular, as it allows for in-depth understanding of complex phenomena and contextualization of research findings within patients' subjective experience (Dattilio, Edwards, & Fishman, 2010; Midgley, Ansaldo & Target, 2014). Previous qualitative outcome studies shed a more pessimistic light on therapy effects (as patients often mention negative therapy experiences despite good outcome scores) and challenge the typical conceptualization of "good" psychotherapy outcome in terms of

measurable symptom or behavioral change (McLeod, 2013). Changes on the level of patients' personality, life, interpersonal relations and self-understanding have for instance been observed (e.g., Binder, Holgersen & Nielsen, 2010; De Smet & Meganck, 2018). Meta-syntheses of qualitative studies on experienced changes by Timulak (2007; 2010) observed changes on the level of awareness, insight, self-understanding, behavioral change, problem solution, emotional experiencing, empowerment and relief (defined as significant events in therapy). Moreover, important similarities across different forms of therapy (e.g., experiences of change on the domains of insight, action and feelings, see Carey et al., 2007) as well as meaningful differences (e.g., applying coping skills in CBT and using self-reflection following PDT, see Nilsson, Svensson, Sandell & Clinton, 2007) have been noted. However, the link between significant events or experienced changes and outcome remains underexplored (Timulak, 2010).

Notably, the few mixed methods studies that focus on the comparison of quantitative and qualitative evaluations of outcome, suggest that the distinction between "good" and "poor" outcome is less clear-cut when examined qualitatively, in comparison to statistically identified categories (McElvaney & Timulak, 2013; McLeod, 2013). These observations may indicate that the groups cannot be so easily separated. Meaning that there is a high need for more elaborative and mixed method research to better understand the relationship between this categorization and the underlying experiences (Atkins et al., 2005). For example, differentiating non-improvement from deterioration is considered necessary as they comprise different phenomena that are most likely caused by different mechanisms (Lambert, 2011). Whether the same applies to the positive side of the outcome continuum - recovery and improvement in psychotherapy - is still unclear. As the field of psychotherapy relies strongly on the JT method for defining the clinical importance of outcome and treatment effects, further exploration of what the statistical classifications of "good outcome" mean in relation to the subjective experiences of the patients who fall within these groups is needed (Atkins et al., 2005).

A recent paper by Rottenberg and colleagues (2018) suggests that a lack of adequate understanding of good outcome might be a particular hurdle in the attempts to understand positive change processes in the case of major depression, generally considered one of the most prevalent mental conditions worldwide (WHO, 2017). The authors more specifically state that the dominant perception of depression as a *chronic* mental illness with poor prognosis has resulted in a "curious neglect of high functioning after psychopathology" (p. 549) and a lack of focus on good outcomes after depression in general. To address this neglect, they argue that successful outcome requires a definition that expands the absence of clinical symptoms (Rottenberg et al., 2018; see also Zimmerman et al., 2012).

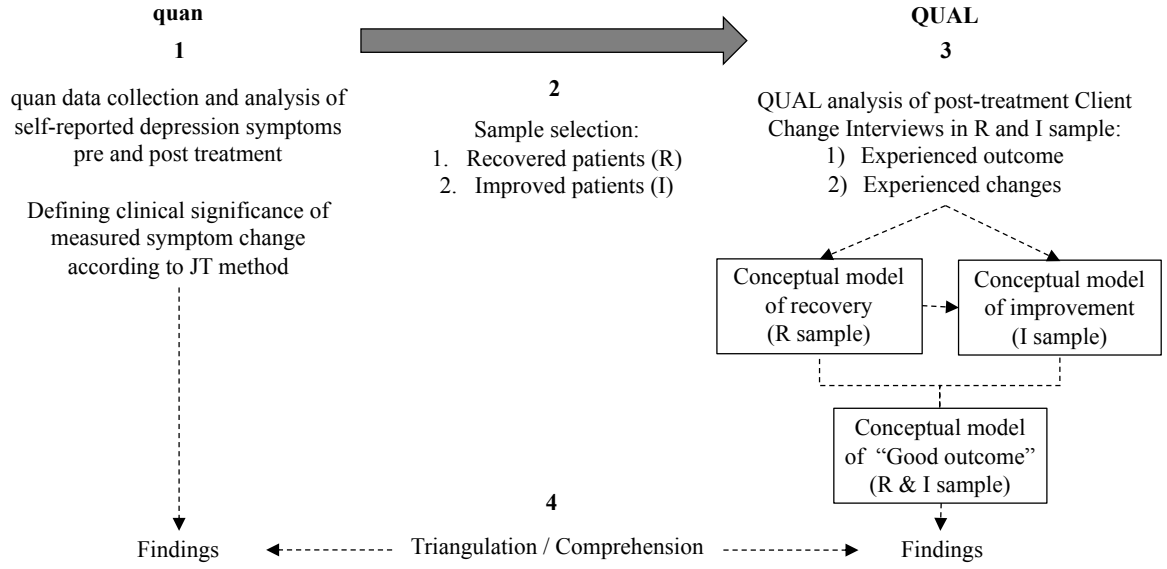
Therefore, in the present study we utilize Major Depression as a representative case to broaden the much-used JT classification of significant outcome with patients' perspectives. In doing so, we investigate whether this type of classification differentiates patients' outcome in a clinically meaningful way, that is, with actual meaning to patients in therapy. The "clinically meaningfulness" is thought as those issues considered as important in patients' experiences. In this study, we explore the experiences of patients who can be categorized as "recovered" and "improved" based on self-reported depression symptoms in order to gain insight into what lays within and beyond the header of "good" psychotherapy outcome. To obtain our objective, the study is structured according to two central interrelated research questions. First, how do patients who recovered and improved in depression symptomatology experience their outcome (i.e., patients' general outcome appraisal)? For instance, do they feel they have improved or recovered themselves? Second, what changes have these patients experienced since the start of therapy? In trying to understand what good outcome means to patients, we address a particular methodological concern in the field of psychotherapy, namely how outcome should be conceptualized and studied in order to provide clinically meaningful outcome findings (McLeod, 2013).

Method

Design

An explanatory sequential mixed methods study (Hesse-Biber, 2010) was conducted, comprising a statistical test of clinical significance in measured symptom changes as well as a qualitative analysis of patients' perspectives. In accordance to the mainstream primary outcome method (see De Los Reyes et al., 2011), a psychometrically sound outcome measure was used to determine patients' outcome (see instruments). Both quantitative and qualitative data were gathered simultaneously; the analyses were done independently and integrated at the interpretation phase (Hesse-Biber, 2010). The samples of recovered and improved patients were studied separately and consecutively, in order to address the statistical differentiation of good outcome as used in psychotherapy outcome research. Figure 1 gives a schematic overview of the applied design.

Figure 1. Explanatory Sequential Mixed Methods Design



Note. Step 4: Qualitative results were interpreted in relation to the statistical outcome classifications, enabling deeper comprehension and triangulation of findings (Hesse-Biber, 2010).

Setting

This study is situated in the Ghent Psychotherapy Study (GPS; Meganck et al., 2017), an RCT on the treatment of major depression (registered on Open Science Framework; ISRCTN

17130982). Patients in this trial were recruited via social media and general practitioners in the area of Ghent, Belgium. The main inclusion criterion was a diagnosis of Major Depressive Disorder (APA, 2013). Further eligibility criteria were sufficient knowledge of the Dutch language and age between 18 and 65. People with a primary diagnosis of substance abuse, acute psychosis and suicidal ideations were excluded. Patients were randomly assigned to time-limited Cognitive Behavioural Therapy (CBT) or Psychodynamic Therapy (PDT). Patients progress was evaluated using questionnaires accompanying every session, interviews were conducted prior to treatment, around the eighth therapy session and after treatment termination. The follow-up period of the study spans two years (ongoing) and consists of four interviews and quantitative assessments. In compliance with APA ethical standards, this study was approved by the Ethical Committee of the University Hospital of Ghent University (Belgium; EC/2015/0085). All participants gave written informed consent prior to participating in the study.

Participants

One-hundred participants were included in the RCT, seventy-four of them provided complete post data and could be considered for our sample selection. For the current study, all patients showing recovery and improvement on the BDI-II (van der Does, 2002) at treatment termination were selected (i.e., purposeful sample and total population sampling strategy). Patients with a pre-score ≤ 11.3 (i.e., clinical cut-off) and patients who dropped out (i.e., patient-initiated premature termination of therapy) within four sessions of treatment (in line with Barrett et al., 2006), were excluded (i.e., n=2). This resulted in a sample of 28 participants who recovered (CBT: n=20; PDT: n=8) and a sample of 19 participants who improved (CBT: n=7; PDT: n=12).

The **recovered sample** consisted of 18 women and ten men, ranging in age between 20 and 60 (M=33; SD=10.8). Two patients were married, eight were cohabiting, sixteen were

single and one was divorced. Sixteen patients had an undergraduate degree, ten patients had a high school diploma. Out of the 28 patients, 17 were employed, four were unemployed and three patients had temporarily interrupted their work due to sick leave. The **improved sample** consisted of twelve women and seven men, ranging in age between 21 and 61 (*M*=40; *SD*=12.6). Three patients were married, seven were cohabiting, six patients were single and two were divorced. At the time of the study, twelve patients had an undergraduate degree, six had a high school diploma. Out of nineteen patients, twelve patients indicated they were employed, one patient was unemployed, and one had interrupted work due to sick leave. In both samples, previous treatment consisted of psychotherapy (for 22 recovered and 16 improved patients) and medication (for nine recovered and six improved patients). All patients had a primary diagnosis of Major Depression Disorder at the start of treatment. Main secondary diagnoses were (as observed in the recovered and improved sample respectively): Anxiety (*n*=6; *n*=5), Phobias (*n*=7; *n*=4), PTSD (*n*=0; *n*=4) and OCD (*n*=1; *n*=4)[18].

Treatment

Treatment consisted of CBT and PDT for major depression, two types of therapy that can be distinguished based on their directive (i.e., CBT) and explorative (i.e., PDT) style of interventions. Therapy was provided by one of four therapists in each condition. Both treatments were manualized and time-limited, consisting of 16-20 sessions. Treatment was delivered with an average frequency of one session per week; sessions lasted approximately 45 minutes. The CBT manual was based on the Cognitive-Behavioral Protocol for Depression by Bockting and Huibers (2011). The PDT manual was based on the Supportive-Expressive Time Limited manual for Major Depressive Disorder by Luborsky (1984; Leichsenring &

[18] Other Axis 1 diagnoses present in 1 or 2 patients in the recovered group included: secondary substance abuse, eating disorder, dysthymia; and in the improved group: eating disorder, secondary substance abuse, somatic symptoms, pain disorder, hypochondria (according to SCID for DSM-IV-TR; First, Spitzer, Gibbon & Williams, 2002).

Schauenburg, 2014). Therapists had an average age of 33 (*SD*=9.6) and had three to eight years of relevant clinical experience and training in CBT or PDT. During the trial, all therapists received two days of training, a patient with which to practice the treatment manual and research procedure under supervision, and bi-weekly supervision sessions.

Instruments

Beck Depression Inventory. The Beck Depression Inventory (BDI-II; Beck, Steer, & Brown, 1996; van der Does, 2002) is a measure of self-reported depression severity. The questionnaire consists of 21 items that are scored on a scale of 0 to 3 and is divided into a cognitive, somatic and affective subscale. A total score between 0-13 indicates minimal depression, 14-19 mild depression, 20-28 moderate depression, 29-63 severe depression. The Dutch version of the questionnaire, BDI-II-NL, shows good validity and reliability (van der Does, 2002)[19].

Semi-structured interview. An adjusted version of the semi-structured Client Change Interview (CCI; Elliott, Slatick & Urman, 2001) was administered. The interview guide was constructed to evoke participants' experiences of therapy, the changes that occurred during therapy, and what they believe influenced these changes, for instance, helping and hindering aspects of therapy. Examples of some questions are: "Which changes have you noticed since the start of therapy (e.g., in relation to others, at school/work, in your emotional wellbeing)?", "How did therapy contribute to these changes?" and "What other factors (outside of therapy) do you think have contributed to these changes?" All interviews were conducted at the end of treatment, approximately one week after therapy ended and took place at the university. Interviews lasted 60 minutes on average. Interviews were audiotaped, and transcripts were analyzed using Nvivo 11 and 12 (QSR International).

[19] A complete overview of the measures used in the GPS can be found in the study protocol of the RCT (Meganck et al., 2017). Given the focus of the current study on major depression, the BDI-II (Beck, Steer, & Brown, 1996; van der Does, 2002) was selected as the outcome measure.

Quantitative Outcome Classification on the BDI-II pre-to-post treatment

Participants were classified in terms of recovery (i.e., clinically significant change) and improvement (i.e., reliable change) based on the Jacobson and Truax (1991) method for outcome classification. Patients self-reported symptom severity was measured prior to therapy and one week after treatment ended. The outcome scores of the patient population were compared to Dutch norms (van der Does, 2002). In order to reach reliable change for the BDI-II total score, a person must show a decrease in scores equal to or larger than 9.6. The cut-off between the clinical and nonclinical population for the Dutch BDI is set at 11.3 (based on the internal consistency of 0.92; van der Does, 2002). This leads to four possible outcomes: Recovery (a decrease in scores ≥ 9.6 and post-treatment score < 11.3), improvement (a decrease in scores ≥ 9.6), no improvement (a decrease or increase in scores < 9.6) and deterioration (an increase in scores ≥ 9.6). Additionally, a post-hoc independent samples t-test was conducted to compare the mean depression pre-scores for the recovered and improved group.

Descriptive Pre-Post Outcome Scores on the Beck Depression Inventory.

Table 1 summarizes the average score on the BDI-II (Beck, Steer, & Brown, 1996; van der Does, 2002) pre- and post-therapy in the recovered and improved sample, presenting findings for the total sample and CBT and PDT subsamples specifically. At the start of treatment, depression severity varied across recovered patients: two patients had a score of mild depression, 12 patients scored as moderately depressed and 14 patients scored as severely depressed. On average, the group mean indicated severe depression prior to treatment. At the end of treatment all patients scored within the non-clinical range, indicating minimal depression severity. The scores of all patients in the improved group indicated severe depression at the start of treatment. At the end of treatment, five patients scored mildly depressed, ten patients scored moderately depressed and three patients scored severely depressed. On average, the group decreased from a mean score indicating severe depression to a mean score indicating

moderate depression. There was a statistically significant difference in scores for the recovered (M = 29.18, SD = 7.42) and improved (M = 38.52, SD = 7.14) group; t (45) = -4.30, p = .00 (two-tailed) at the start of treatment, indicating the improved sample scored significantly higher prior to therapy. The magnitude of the differences in the means (mean difference = -9.35, 95% *CI:* -13.72 to -4.97) was large (Cohen's d = 1.28).

Table 1. Pre-Post Outcome Scores on the BDI-II: Recovered and Improved Sample

	Recovered sample						**Improved sample**					
	Total (*n*=28)		CBT (*n*=20)		PDT (*n*=8)		Total (*n*=19)		CBT (*n*=7)		PDT (*n*=12)	
Total scores	M	SD (range)	M	SD (range)	M	SD (range)	M	SD (range)	M	SD (range)	M	SD (range)
Start therapy	29.2	7.4 (16-46)	29.7	7.2 (16-46)	28	7.37 (20-38)	38.5	7 (30-53)	36.3	4.6 (30-45)	39.8	7.6 (30-53)
End therapy	4.7	3 (0-9)	4.5	2.8 (0-9)	5.1	3.30 (0-9)	22.4	7.9 (13-41)	19.7	4.4 (13-22)	24	8.8 (14-36)

Note: Total scores on the BDI-II (van der Does, 2002): 0-13: minimal depression, 14-19: mild depression, 20-28: moderate depression, 29-63 severe depression. RCI: ≥ 9.6; clinical cut-off: < 11.3.

Qualitative analysis: Grounded Theory

The explorative and interpretative method of grounded theory was selected as method of choice as it allows for the construction of new rationales grounded in data (in our case patient interviews), which corresponds to our aim of understanding patients' experienced changes within and beyond the standard outcome classification (Charmaz, 2014; Glaser & Strauss, 1987). The first author in dialogue with the third author conducted the analyses; the second author functioned as an auditor throughout the process (Hill et al., 2005; see below). The first author selected the samples and interviews corresponding the sequential explanatory design outlined above (see Figure 1). The second and third author were aware of the focus of the study being good outcome cases and the existence of a recovered and improved sample of patients

but, were unaware of patients' specific outcome classification (i.e., whether patients belonged to the recovered or improved sample as well as the actual scores on the BDI-II) during the open coding phase of the data analysis. This was done to ensure that the researchers' perspectives remained open to findings contrasting potential pre-existing expectations. From the axial coding phase on (see below), all researchers were conscious of which themes belonged to the recovered or improved group.

Pertaining to the characteristics of grounded theory, several stages of analysis were completed in a cyclic manner before arriving at the final results (Mortelmans, 2013). The analyses were first conducted with the interviews of the recovered sample and the interviews of the improved group were analysed thereafter. During the analyses, existing categories were reused if substantial similarities were observed between the groups. However, where appropriate, categories were adjusted, or new ones were created to ensure that particular experiences in both samples were not overlooked. In the final stage of the analyses central themes grasping the experiences of both samples were defined (see Figure 1).

During the *open coding phase*, the interviews were extensively read to identify relevant parts (i.e., meaning units; Giorgi & Giorgi, 2003) of the interviews relating to the research questions. Labels were attached to certain parts of the text, differentiating patients' experienced outcome and experienced changes. Non-relevant parts, i.e., not dealing with the topic of well-being or experienced changes were omitted from the analysis. This phase resulted in an initial list of codes formulated with the intent to remain close to the narrative of patients. The list of codes was organized thematically in the *axial coding phase*. For the recovered sample, this resulted in the themes: "empowered self", "altered interpersonal relationships", "calmer self" and "increased understanding". For the improved sample, the themes: "growing self", "relational issues", "insight/voice of the therapist", "letting things go" and "ongoing difficulties" were extracted. The *selective coding phase*, during which the theory was refined

by assembling the themes in core domains of change, resulted in “empowered self” and “strengthened but struggling self” as core categories for the recovered and improved sample respectively. The aforementioned sample-specific subthemes were constructed around these core categories aiming to grasp the experiences expressed in each sample. Ultimately, the resulting core and subthemes in both samples were re-examined in order to formulate overarching core domains and subdomains of change that could grasp the experiences of both the recovered and improved patients. This allowed answering the question, which experiences lay within and beyond the header of “good outcome” while at the same time highlighting both significant similarities and differences across both samples. The salience of the resulting domains and subthemes were indicated by means of the nomenclature proposed by Hill et al. (2005) and criteria for larger samples by Knox et al. (2006): General: ≥ 90% of the participants (25–28 and 17-19 recovered and improved patients respectively); Typical: ≥ 50% and < 90% (14-24 and 10-16 patients); Variant: ≥ 20% and < 50% (6-13 and 4–9 patients); Rare: < 20% (1-5 and 1-3 patients).

In an answer to the first research question, patients’ general outcome appraisal (e.g., “I’m doing better”, “today I feel better, tomorrow we’ll see”) were presented descriptively (e.g., as indicating “improvement”, “ups and downs”). Answering the second research question, we summarized patients’ experienced changes (see Table 2) and constructed a conceptual model (Figure 2), which depicts the explicit purpose and findings of this study: elucidating the meaning of “good outcome” beyond the statistical classifications of “recovery” and “improvement”. The results were outlined using illustrative quotes (recovered patients were indicated by numbers 1 to 28 and improved patients were given the letters A to S to guarantee anonymity).

The research team conducting the analyses was gathered aiming at a balance in terms of familiarity with the interviews, academic and clinical training background and phase in

career, to ensure a horizontal and open dialogue (see Parker 2007). However, we acknowledge the inevitable influence resulting from the perspective of the researchers conducting the qualitative analyses. The first author and primary analyst in this study, is a white female of 28 years old. She is an early career researcher, trained in qualitative, mixed methods and quantitative research, her main expertise is process-outcome research in psychotherapy, depression and ADHD. She focuses on patients' lived experiences and adheres a critical stance towards potential contextual influences on patients' self-understanding. She is a clinical psychologist and has three years of clinical experience, mainly in working with adolescents in inpatient and outpatient facilities using a solution focused and psychodynamic approach. She was involved in the data-gathering process and conducted part of the interviews that were used in this study. The second author and auditor during the analysis, is a white female of 37 years old. She is an associate professor and an experienced researcher trained in quantitative, mixed methods and qualitative research. She is a clinical psychologist and psychotherapist trained in psychoanalytic psychotherapy. She has nine years of clinical experience in private clinical practice, mainly in adult outpatient care. She was involved in the conception and development of the design of this study and the Ghent Psychotherapy Study. The third author who aided in the analysis in this study, is a white female of 28 years old. She is an early career researcher, trained in qualitative and quantitative research. She is a clinical psychologist and psychotherapist trained in psychoanalytic psychotherapy. She has six years of clinical experience in inpatient and outpatient clinical practice (both psychodynamic and solution-focused therapy), with adolescents and adults. She was involved in the data-gathering process, conducting part of the interviews that were used in this study. As clinical psychologists and psychotherapy researchers, all analysts have a background in multiple theoretical and methodological orientations, which reflects their open stance towards different forms of psychotherapy and methodologies.

Potential biases resulting from the analysts' academic, theoretical and clinical background, were bracketed as much as possible throughout the process of the study and conducting the analysis. First, the main affinity of the analysts is psychodynamic theory, which potentially influenced the way patients' experiences were interpreted. This influence was countered as much as possible, by taking the patients' perspectives as the main focus of analysis. Rather than a theoretical interpretation of their narratives, we stayed close to patients' own words (i.e., using their own descriptions), using a grounded theory bottom-up approach. Second, the analysts share a personal interest in idiosyncratic experiences and departed from the assumption that psychotherapy is a complex phenomenon that should be studied accordingly. This clearly navigated the focus of the study and mixed methods analysis, however, by making this explicit, consequences of implicit guiding assumptions were precluded. It could be expected that an idiographic stance resulted into individual differences being more easily noticed than the drawing of generalized conclusions. We acknowledge this influence yet believe this is also the value of the qualitative analysis that was conducted. This was moreover countered as much as possible by conducting the analyses in several phases, trying different ways of presenting and describing the findings until a balanced representation of the findings was found. Finally, the first analyst was aware of patients' outcome scores and was most familiarized with patients' data, which inevitably guided her interpretations of patients' experiences. The second and third analyst challenged this influence, as they were explicitly unaware of patients' outcome scores during the initial phases of the analyses (i.e. triangulation among researchers and interviews was applied).

Credibility checks were held at several stages of the analysis. To ensure no relevant information was overlooked in the interview, every patient was asked whether they wanted to add further information. During the study preconceptions were left aside (i.e., "bracketing") as far as possible, by means of three main strategies that were applied at several stages of the

analysis. First, while conducting the interviews, we applied "focused interviewing" strategies in order to elucidate patients' experiences (i.e., using semi-structured interviews with open explorative questions) rather than own preconceptions. Second, notes and ideas were written down in a "reflective diary" and by using memo's (e.g., in Nvivo) by the first author throughout the entire analyses. When discussing the themes resulting from the interviews, these preliminary ideas were discussed with the third and second author. Also, the first author wrote a short summary for every patient prior to coding the interviews, in order to grasp every individual narrative allowing to go back and forth between themes that were generated for the entire sample of patients and the individual stories (i.e., checking whether the themes still aligned with patients' stories). Lastly, the second author participated as an "independent researcher" (i.e., who was not involved in conducting and analysing the interviews) and audited the process to actively test assumptions. For instance, perceived differences and similarities between the samples were challenged using critical questions, stimulating the first and third author to rethink or reformulate themes as well as the conclusions that were drawn. In contrast to initial expectations, the differences between recovered and improved patients and PDT and CBT patients appeared smaller than presumed, for instance. Generally, a contextual perspective led the analyses, departing from the assumption that the broader social context influences how patients give meaning to their experiences (Boyatzis, 1998).

Results

"Good outcome" Contains Experiences of Empowerment, Finding Personal Balance and Ongoing Struggle

The findings of the grounded theory analysis are visually presented in the conceptual model depicted in Figure 2; Table 2 offers a more detailed summary. First, we outline recovered and improved patients' general outcome appraisal as reported after ending psychotherapy. Thereafter, we describe recovered and improved patients' experienced changes, which were

grasped by three main domains: *Feeling empowered, Finding personal balance* and *Ongoing struggle*. All three domains are further divided into one or more subdomains, as summarized in Table 2 and described in more detail below. The salience of the domains is indicated with "general", "typical", "variant" or "rare" depending on the number of cases to whom the domains applied (Knox et al., 2006;). These descriptions allow indicating both cross-sample similarities and within-sample particularities. When describing the domains and subdomains we refer to the recovered and improved sample in general (i.e., regardless of the therapy condition). If notable differences were observed for patients in CBT and PDT this is mentioned explicitly.

Experiencing Outcome: Improvement *and/or* Remaining Difficulties Indicate an Ongoing Process and Variety in Experiences. After ending therapy, patients in both samples typically described a certain degree of *improvement* (e.g., "feeling very good", "better") indicating a more constant mood (e.g., "not feeling down anymore"), or an alleviation of distress (e.g., "it feels lighter now"). *Remaining difficulties* (see also Ongoing struggle), were however generally expressed by both groups of patients. Patients in both samples typically described their outcome as *an ongoing process*. Not all difficulties had been altered, some changes required more time or patients still had lingering questions they needed answers to. For instance, two patients who were classified as "recovered" expressed that their core problem remained unresolved.

> 'A lot got untied in therapy, and some of those things are still loose ends, like not all pieces of the puzzle are put together yet.' (Patient 9; PDT; Recovered).

Experiences of *ups and downs* at the end of therapy were typically mentioned by improved patients (e.g., "today I feel good, but we'll see what tomorrow brings") and rarely by patients in the recovered sample. Moreover, eight patients who were classified as "improved" explicitly stated they did *not* feel better (of which one patient felt worse) or stated they had experienced

a relapse after a temporary improvement. Overall, a more unstable improvement and larger heterogeneity in patients' outcome experiences was observed in the sample of patients who were classified as "improved".

> 'Gosh, I'm not going to say that, you know, after those twenty sessions, I cannot say it's going any better, but the therapist always claims that by talking and talking and patience and patience we will get there… so, I'm giving her the benefit of the doubt and shall continue therapy.' (Patient J; PDT; Improved).

Accordingly, the therapy was typically considered "sufficient" for recovered patients, (i.e., they could manage on their own) as well as for a part of the improved patients (i.e., mentioned variantly). Improved patients typically considered further therapy important to sustain the acquired changes, which was only rarely expressed by recovered patients.

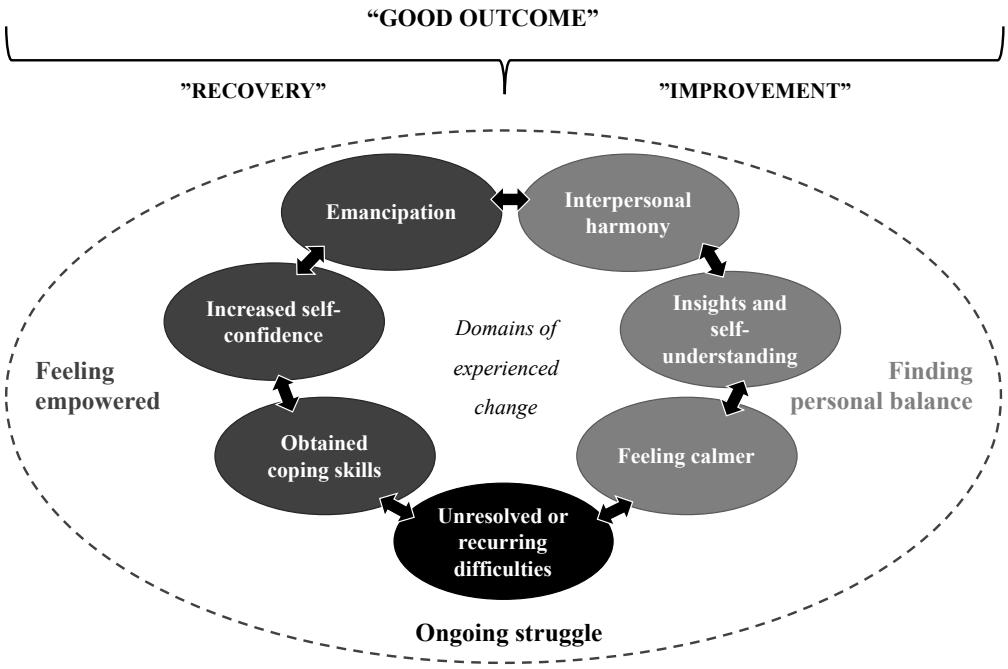
> 'I do feel that I will remain susceptible [to depression] for a while, so after the final session with the therapist I asked if she knew someone to continue my treatment. I see the necessity of that. I think I can use some more therapy to prevent me from relapse.' (Patient 10; CBT; Recovered).

> 'I am still often scared I will get back into it [the depression] and sometimes I am not doing so well, I would like to make some more progress in that in order to become a bit stronger in general.' (Patient I; PDT; Improved).

Nonetheless, in both samples several patients considered therapy insufficient on its own, or uncertain in effect. That is, recovered and improved patients variantly expressed certain degrees of *disappointment* (e.g., the sessions were too short, too repetitive, not enough to facilitate change).

Feeling Empowered, Finding Personal Balance and Ongoing Struggle. Aside from the general post-treatment appraisal of outcome, patients mentioned experienced changes within the domains of *empowerment*, *finding personal balance* and *ongoing struggle* (see Table 2 and Figure 2). Here, we describe and illustrate the subdomains and how they were expressed by patients in the recovered and improved sample.

Figure 2. Conceptual Model of "Good Outcome"



Note. The dotted line indicates the interrelatedness of domains. The bidirectional arrows assume a mutual influence between the subdomains.

Table 2. Taxonomy of Changes Experienced by Recovered and Improved Patients

Domains of experienced change	Recovered	Improved	*
Feeling empowered	**General**	**Typical**	*
Increased self-confidence	*Typical*	*Typical*	
Speaking up and being heard	Typical	Typical	
(Becoming) who I want to be	Typical	Variant	*
Getting rid of "mask of positivity"	Typical	Rare	*
Emancipation	*Typical*	*Typical*	
Healthy self-centeredness	Typical	Typical	
Going my own way	Typical	Variant	*
Obtaining new coping skills	*Typical*	*Typical*	
Controlling thoughts and canalized emotions	Typical	Typical	
Tackling fears and challenges	Typical	Rare	*
Remaining issues but different perspective	Variant	Typical	*
Finding personal balance	**General**	**Typical**	*
Interpersonal harmony	*Typical*	*Typical*	
More positive relationships/interactions	Typical	Typical	
Separation (from negative relationships)	Variant	Typical	*
Insights and self-understanding	*Typical*	*Typical*	
Insight into past, self and problems	Typical	Variant	*
The echo of the therapist	Rare	Variant	*
Feeling calmer	*Typical*	*Typical*	
Able to let go and mindful stance in life	Typical	Typical	
Milder towards self and (process of) acceptance	Variant	Variant	
Ongoing struggle	**General**	**General**	
Unresolved or recurring problems	*Typical*	*Typical*	
Inability to move forward (lack of energy)	Rare	Typical	*
Vulnerability and ambivalence regarding change	Rare	Typical	*

Note. *: highlights the differences between the recovered and improved sample in terms of the salience of the domains (according to Knox et al., 2006).

Feeling empowered. Across both samples of patients, three domains of change were differentiated that indicate feelings of *empowerment*. Patients in both samples typically mentioned *increased self-confidence, emancipation* and *obtaining new coping skills*.

Increased self-confidence. Typically for both recovered and improved patients, self-confidence was expressed in terms of an increased self-esteem or personal growth; especially for patients in the improved sample this was often still on-going, e.g., "I feel that I am developing my talents now". Patients felt better about themselves and were more aware of what they are worth.

> 'Now I think: no-one is more or better than me. I am equal to others. I always had the tendency to think people around me were smarter, better, to make myself inferior.' (Patient 7; PDT; Recovered).

> 'I became aware of the fact that, yeah, you know, that I'm still capable of a lot of things. I'm aware that I actually do have certain talents and gifts and I dare to use them again more spontaneously.' (Patient Q; CBT; Improved).

Typically mentioned by patients in the recovered sample, but only variantly by improved patients, was that they now *dare to speak their mind* to others, dare to indicate what they want and consequently they felt they were listened to.

> 'I really claimed my spot, I'm now a worthy member of the team. If something needs to be done, I now dare to ask someone else to do it.' (Patient 15; CBT; Recovered).

Moreover, the recovered sample typically mentioned they felt empowered *to get rid of their "mask of positivity"*: patients no longer hid their negative emotions or weaknesses and consequently also felt that they stayed truer to themselves, indicating an increased openness.

> 'My cheerful moments feel genuine now. Before I often played to be happy to hide my sadness. But now, if I feel sadness I can seclude myself for a while, but not uhm. It's sincere happiness, it's less exuberant as before, but you know, before it was fake, I would act crazy, play the "crying clown".' (Patient 1; PDT; Recovered)

Emancipation. Empowerment was also reflected in the fact that patients in both samples typically indicated they took more care of themselves by respecting their personal boundaries and putting themselves first. While before, other people had a very dominant impact on their lives and patients gave more to others than they did for themselves, after therapy this changed markedly (see Interpersonal harmony). For instance, patients said they cared less about other people's problems, expressing *healthy self-centeredness.*

> 'I want to be myself and I really will not put any more energy into people who don't have good intentions. That's my motto for the moment: I'm not investing in things that will gain me nothing. I don't think that's selfish, but more like healthy selfishness. It means considering yourself as well.' (Patient A; PDT; Improved)

Emancipation was also reflected in finding an *own voice* and making *own life choices*, typically mentioned in the recovered sample and variantly in the improved sample. Instead of considering other people's opinions, patients were more certain about who they were and what they wanted in life. Patients stated they were able *to go their own way,* regardless of "what is expected of them".

> 'The fact that, you're expected to, when you graduate, to look for a job and find a job immediately and throw yourself at it completely. I mean, I think I've learned not to listen to that and to look at myself first. You know, what do *I* want? Rather than starting to work right away.' (Patient 2, PDT; Recovered)

In the improved sample, this subdomain was generally found for patients in CBT and typically for patients in PDT.

Obtaining new coping skills. A final subtheme illustrating patients' empowerment, is the fact that patients gained new *coping skills*. In both samples this was typically described as having grip on thoughts and emotions. In comparison to the beginning of therapy, patients described they had learned to get a hold of their thoughts, e.g., being able to stop negative thinking. In the improved group, this was generally mentioned by patients in CBT (e.g.,

referring to concrete tools to control thoughts) and typically by patients in PDT. Also, emotions were more canalized than before (i.e., less overwhelming; mentioned by patients in CBT and PDT). Recovered patients moreover typically said their fear for the future had diminished, as they felt armed to *tackle future challenges*; they regained energy and motivation to go on, which was not the case in the improved sample (see Ongoing struggle).

> 'I feel that I'm armed now, that I can handle misfortune better because I've gained more insight into myself and have ideas about how to handle certain things.' (Patient 27; CBT; Recovered)

Typically for improved patients, on the other hand, was the fact that certain situations or issues did not alter, but that they now tried to see them in a *different perspective* or deal with them differently.

> 'At work there are a lot of changes going on, and normally I would really try to sabotage them and fight against it but yeah I now just try to accept it' (Patient C; PDT; Improved)

In line with these findings, recovered patients stated variantly that they considered remaining problems as "normal", for instance mood swings became signs of not being "lifeless" and "off-day's" can happen.

In sum, the first main domain expresses patients' empowerment, reflected in the fact that they felt more confident about themselves, which coincided with the ability to set their own course independent from others. Contributing to this, patients also obtained new coping skills, which at the same time expressed their increased self-confidence.

Finding personal balance. Interrelated with increased empowerment, patients achieved, or were in the process of finding, a personal balance. This was typically reflected in *interpersonal harmony, insights and self-understanding* and *feeling calmer.* Notably, in both the

recovered and improved sample this domain was more often mentioned by patients following PDT (generally) than in CBT (typically)[20].

Interpersonal harmony. First, patients' emancipation (see Feeling empowered) was reflected in how they related to others or how relationships were experienced. Typically for both samples of patients, more positive relationships or interactions were experienced. Negative relationships or interactions diminished while more positive interactions were cultivated: increased closeness to significant others was mentioned, as well as different perspectives on certain relationships (e.g., understanding protective behaviour of parents better). Patients also engaged more easily in social interactions.

> 'I don't need to go out and have fun outside of my marriage anymore. (…) I can enjoy things more with my wife, I don't need to go out that much anymore to have my freedom, we [patient and his wife] do more things together again.' (Patient 26; CBT; Recovered).

However, a tendency to take distance from others (*separation*) was described, typically by improved patients and variantly in the recovered sample. Improved patients typically stated they experienced certain people as intrusive or considered them "negative influences" who they now actively tried to avoid (see Healthy self-centeredness).

> 'There is one friend I don't talk to anymore. She was so pushy, so demanding. I still think of her sometimes, but I'm so happy I don't have the pressure anymore to always do good for her' (Patient E; PDT; Improved).

Insights and self-understanding. Finding a personal balance was stimulated by an increased self-understanding or gained insights, typically mentioned in both groups of patients. Especially in the recovered sample, patients said they could understand themselves better and could now see what had caused their difficulties (e.g., the important influence of their past). Furthermore,

[20] In the recovered sample this was the case for all three subdomains, in the approved sample for two out of three subdomains ("feeling calmer "was typically expressed by patients in CBT and PDT).

they recognized problematic aspects in the way they lived their lives and realized the impact they themselves had on their problems.

> 'It was extremely confronting, I laid all responsibility on my parents or the situation in my family but yeah I was really shocked to realize what I had done myself to sustain that situation.' (Patient 5; PDT; Recovered)

In the improved group, similar insights were mentioned variantly, but in contrast to recovered patients, equally often presented as *an echo of the therapist* (e.g., "the therapist says I keep things inside, I'm a time bomb") and most often by patients in PDT (i.e., typically versus variantly in CBT).

> 'I know I should listen to my body, you know, the therapist said I should really listen to my body more.' (Patient N; CBT; Improved)

The fact that patients echoed the words of the therapist indicates the therapists' opinion was highly valued and show these insights were important and strengthening, yet seemingly they were not yet or not entirely internalized.

Feeling calmer. Coinciding with an increased self-understanding and interpersonal harmony, patients in both samples felt "calmer". This subdomain captures patients' typical experience of being able to *let things go or see things in perspective*. Improved patients for instance indicated they enjoyed things more and could live more in the "here and now". Patients in the recovered sample typically expressed a more *mindful* or conscious way of living. Instead of rushing things, patients learned to take their time, which was described as a more "slowed down" stance in life. Variantly stated in both samples, patients had also become *milder* towards themselves; they had lowered their standards and expectations and accepted personal limitations and mistakes (see Getting rid of "mask of positivity").

> 'I'm less strict with myself and in the end that makes life so much easier. I used to draw very negative conclusions about myself when I did something wrong and I considered it to be part of my personality, that I wasn't doing things right. And now I allow myself to make more mistakes.' (Patient 3; PDT; Recovered).

However, in the improved sample a *process of acceptance* was described as still ongoing, for instance because of ongoing external stressors (e.g., divorce or chronic illness), and sometimes coinciding with lasting signs of distress (e.g., nightmares, headaches, irritability).

> 'When you are chronically ill you go through a certain mourning process and my process is not over yet, I still cry a lot, that's not bad. (…) It's a matter of accepting it.' (Patient S; CBT; Improved)

Taken together, the second main domain captures how patients experienced (increased or being in the process of finding) personal balance through improved interpersonal relationships, gaining insights and self-understanding and a general sense of feeling calmer after therapy.

Ongoing struggle. Despite patients' experienced empowerment and finding a personal balance, ongoing struggles were mentioned generally in both samples. Typically, this meant that certain issues remained *unresolved or reoccurred* (see Remaining difficulties, supra).

> 'I still think I don't have a normal way of functioning, I mean, if I have to work so hard to stay healthy, that's not normal, that's just not… (…). I really have to put so much effort to stay stable, to function normally (…). I am in a better condition than I was in the beginning, but I still have irrational thoughts and fears and they keep coming back.' (Patient 12; CBT; Recovered)

Expressions of ongoing struggle were predominantly observed in the improved sample and within the sample generally mentioned by patients in PDT and typically stated by patients in CBT. Specifically, patients typically described *an inability to move forward*, particularly due to a lack of energy or motivation (e.g., "the work just lays there, it piles up, I know I should get it done, but just can't get to it."); the recovered group in contrast did feel re-energized to tackle future challenges (see Obtained coping skills, supra). Moreover, improved patients typically said they felt as if one negative event could sweep them of their feet again, expressing *enduring vulnerability* or unstable improvement. Also, they expressed *ambivalence* about the changes they were undergoing, as they did not feel entirely natural or they were afraid to hurt others, for

instance, when standing up for oneself could offend others (see Emancipation; Interpersonal harmony).

> 'I would like to say no to people more easily. I still don't feel good about it, although I do say it now. It's still not like "I say no and I'm okay with it". I'm upset for a few days.' (Patient M; PDT; Improved).

As such, in contrast to the recovered sample, it seems that the increased empowerment and personal balance were less self-evident for part of the patients who were classified as "improved", in line with the heterogeneity in outcome experiences in this sample.

Discussion

This study examined the subjective experiences of patients whose outcome indicated recovery and improvement in depression symptoms following psychotherapy in the context of a randomized controlled study. Doing so, we elucidated what "good outcome" in psychotherapy research means in light of patients' own experiences. In the following, we discuss whether the influential JT framework of clinical significance offers a meaningful differentiation of patients' functioning after therapy and elaborate on how good outcome can be understood from patients' perspectives.

Firstly, the findings of this study showed that "good outcome" based on pre-to-post symptom evaluation contained patients' experiences of improvement *and/or* remaining issues, indicating an ongoing process and variety in experiences. Thus, patients in this study offer a more nuanced or pessimistic perspective on psychotherapy outcome findings, corresponding to what has been observed in previous qualitative outcome studies (see McLeod, 2013). Recovered patients did not call themselves "cured" after therapy; rather, they mentioned significant changes *despite* residual difficulties, implying a need for further therapy for some. The improved sample was specifically characterized by heterogeneity in experiences (ranging from feeling better to not doing well at all). Correspondingly, the self-reported symptom severity in this group varied widely (from severe to mild depression) at treatment termination.

As such, not all patients agreed with the classification of "good outcome". For instance, some recovered patients, indicated core difficulties had not been altered. In the improved sample this discrepancy was more outspoken, as for some, an ongoing struggle dominated their experience post therapy. Almost half of the patients in this category reported that they did not experience improvement, thus contradicting how their outcome was labelled. Consequently, while the JT method of clinical significance can indicate improvement in general terms, it seems to fail to distinguish clinically relevant variations. In line with findings by Zimmerman et al. (2012), reliable and clinically significant decrease in symptoms neglected important domains of remaining difficulties according to patients' experiences[21]. Moreover, the heterogeneity amongst the improved patients suggests that grouping these people within one category is unjustified, as the variation in their experiences asks for very diverse therapeutic proceedings that may easily be overlooked under the header "improvement".

Beyond patients' general outcome appraisal, three main domains of experienced changes were found across both samples, indicating that patients who recovered and improved in symptom severity felt a certain degree of empowerment, personal balance and ongoing struggle. The first main domain of experienced change, empowerment in psychotherapy, is a recurrent finding in qualitative outcome studies (Timulak, 2007; 2010) across different populations and settings (e.g., multicultural psychotherapy, Delgado-Romero et al., 2016). Psychological empowerment is defined as perceptions of personal control, a proactive approach to life, and a critical understanding of one's environment (see Zimmerman, 1995). Timulak (2007) refers to empowerment in terms of personal strength to cope with problems, self-perceptions, interpersonal difficulties and having a sense of personal development. In our study,

[21] Notable, patients in this study used "depression terminology" (i.e., specific descriptive complaints as summarized in the DSM or BDI, e.g., changes in sleep pattern) only minimally or superficially to describe their changes in therapy (e.g., "I feel less down"). A detailed discussion is beyond the scope of the current study, however, further research into the relationship between of BDI or DSM domains and patients' experiences with recovery from depression would be highly informative (see Ratcliffe, 2014).

empowerment was reflected in an increased feeling of self-confidence, emancipation and gaining new coping skills (similar to Ekroll & Rønnestad, 2016; Nilsson et al., 2007). Differences between recovered and improved patients mainly indicated that openness (e.g., getting rid of a mask of positivity), being who they wanted to be and going their own way regardless of others' opinion as well as feeling empowered to tackle future challenges was less typical of improved patients' experiences (or still in development).

The second main domain yields that patients experienced increased personal balance, defined in terms of gained insights and self-understanding, interpersonal harmony and feeling calmer. This observation shows resemblances to general relief experiences described in previous studies (see Timulak, 2007; 2010). Patients' sense of empowerment seemingly coincided with improved interpersonal interactions, ultimately resulting into better interpersonal relationships. This is in line with what has been observed elsewhere, for example by Binder et al. (2010) who found that intra-personal changes coincide with inter-personal changes. For the participants in our study, self-confidence and emancipation in that sense seemingly aided the separation from others and/or reaching a more ideal distance-closeness towards others. Differences between the samples suggested this balance was still being searched for in the case of several patients who were classified as "improved", suggesting the experience or presence of more negative interpersonal relationships. Both recovered and improved patients could however perceive relationships in a different light, reflecting their gained insights and increased self-understanding.

In psychotherapy literature, insight is indeed considered to be an important step in gaining positive outcome. In our study, these insights mainly regarded understanding oneself, one's problems, the influence of past experiences and realizing own implications in current problems (see Nilsson et al., 2007). Patients' increased self-confidence, emancipation and coping skills in this study support the hypothesis that agency operates as an important

mechanism in this relationship, in which agency is seen as the ability to translate insight into concrete action (see Hill, 2004, for a stage model of therapy). This could suggest the hypothesis that the overall differences among the recovered and improved sample depend on differences in agency. In the improved sample, changes were seemingly more precarious and still largely dependent on others (e.g., insight in terms of the echo of the therapist, suggesting a more passive patient position and important role for the therapist), which might illustrate a less agentic position of these patients in general (see Bohart & Wade, 2013, on the role of agency, and Messer & McWilliams, 2007, on levels of insight). Further research into the relationship between outcome and significant events in therapy (e.g., insight) and the mechanisms facilitating change is however required to answer these questions (Carey et al., 2006; Timulak, 2010).

The third main domain, the ongoing struggle that was found for both the recovered and improved sample, is remarkable as it bears resemblance to non-improved patients' experiences, typically classified under the header of "negative outcome" (see De Smet et al., 2019; Werbart, von Below, Brun, & Gunnarsdottir, 2015). Good outcome did not mean all problems were solved, which is in line with the observation that non-improvement in outcome scores did not indicate "no change" from patients' experiences. Recovered and improved outcome resembled an ongoing process for the majority of patients, in line with previous observations (see review by Gianakis & Carey, 2008). Notably, in both samples, remaining issues were dealt with differently or considered normal, which is in line with a more positive perspective on recovery and health in general (i.e., focus on possibilities and abilities, *given* psychological, physical and social challenges; Slade & Longden, 2015). In contrast to patients who did not improve, the recovered and improved patients clearly described a more established feeling of empowerment, insight, personal balance and better interpersonal relationships. However, the difficulties mentioned by the improved sample show striking parallels with what was summarized as a

feeling of being stuck between "knowing versus doing" in a non-improved sample (see De Smet et al., 2019), namely the inability to move forward. Enduring vulnerability and doubt regarding ongoing changes on the other hand seems typical for the improved sample. Nonetheless, the similarities suggest that in light of patients' experiences, the distinction between no reliable change and reliable improvement may be arbitrary.

The comparison of patients' experiences depending on the type of therapy was not the main focus of this study, however, some observations are notable in resemblance to studies that found important differences between CBT and PDT (e.g., Nilsson et al., 2007) and studies emphasizing the similarities between changes experienced in different types of therapy (Carey et al., 2007). Also in this study, we observed important similarities in the experienced domains of change across both types of therapy. In line with the study of Nilsson et al. (2007) PDT patients indicated a feeling of understanding oneself better than before, and likewise this was reflected in altered interpersonal relationships (setting own limits, less intrusive relationships), which was more often mentioned by patients engaging in PDT than by CBT patients, both for the recovered and improved patients. Regarding the CBT condition, similarities with findings of Nilsson et al. (2007) were observed mainly regarding patients obtained coping skills, especially for the improved patients. Ongoing struggle (in terms of vulnerability and ambivalence) was more often mentioned by improved patients in PDT in line with findings by Nilsson et al. (2007). They observed that, despite improvement in both samples, treatment results were most clear to patients in CBT while PDT patients expressed stronger feelings of ambivalence. This suggests different processes of change and patients' experiences depending on the more directive and explorative nature of the treatments. Further research into the mechanisms underlying different types of therapy, for instance the function of insight, could mean an important contribution to these preliminary, yet clinically relevant observations (see also Timulak, 2010).

Several consequences can be drawn from our findings addressing the conceptualization of clinically significant outcome and evaluation of change in psychotherapy research. First, the observation of both similarities and differences between the recovered and improved sample provides support for an outcome continuum (e.g., Tingey et al., 1996) rather than a dichotomy of functional/dysfunctional distributions that underlies the JT conceptualization[22]. One of the main critiques on the JT classification is the unreliability of the differentiated categories at the category ends, as the distinctions between the different groups seem less clear-cut than intended (Lambert & Ogles, 2009). Nonetheless, our findings suggest that a symptom-based distinction may not always be meaningful in light of patients' experiences, and consequently, a continuum based on outcome would still not do justice to the variety of experiences that do yield varying clinical implications. In that sense, good outcome in psychotherapy for depression seems to require a multi-dimensional understanding, in line with the complex and diverse experiences of depression itself (Ratcliffe, 2014; Rottenberg, 2018).

Moreover, our findings raise the question whether the patients who experience an ongoing struggle in terms of enduring vulnerability and inability to move forward (typically patients from the improved sample) are merely "not there yet", which would yield the assumption of a linear pathway towards a similar outcome for all patients. As suggested by several authors, some patients may require more time to gain considerable improvement, which generally argues against the use of uniform time limits for all patients (see De Geest & Meganck, in press, for a review on the impact of time limits in therapy). This hypothesis is

[22] Clarification: Jacobson and Truax (1991) defined these distributions in terms of a dysfunctional and functional *range*, and thus assuming an outcome *continuum.* Nonetheless they assumed "that both functional and dysfunctional people form distinct distributions on the target problem" and thus "clinically significant change can be thought of as movement on the part of an individual from the dysfunctional to the functional distribution." (see Jacobson, Folette & Revenstrof, 1984, p. 340). The "dichotomy" in our understanding thus results from the assumption that distinct distributions exist based on normative cut-off points (one for the functional and one for the dysfunctional population). This has been criticized by suggesting that, for some problems, the idea of one population containing both functional and dysfunctional individual is more likely (see Wampold & Jenson, 1986); for instance, Tingey et al. (1996) developed this continuum-idea by suggesting the use of multiple cut-offs. For a discussion see Wise (2004).

supported by the larger number of patients in the improved group who mentioned that further therapy was needed to increase their strength and facilitate further change. Also, patients in the improved sample scored significantly higher than recovered patients prior to the start of therapy, which could indicate a higher degree of patients' disturbance typically associated with complicating the therapy process and patients' improvement (Mohr, 1995). Nonetheless, a linear idea of change is generally being challenged (see Evans & Fletcher, 2013), and we agree with Carey et al. (2006) that research on patients' change has gained too little attention; as such, its conceptualization requires more detailed experiential consideration to answer this question. Process research that looks into moderators (i.e., for whom and under what conditions does change occur, which can be different depending on patient or therapist characteristics) and mediators of change (i.e., why change occurred, e.g., suggesting the influence of context variables for instance) is required (see Evans & Fletcher, 2013).

Finally, with regard to the attributed effects of therapy, our findings show that in both groups, the experience of outcome did not necessarily correspond to the experience of therapy or further need for treatment, which is in line with previous observations (e.g., De Smet et al., 2019; Werbart et al., 2015). Hence, "good outcome" as defined based on the statistical interpretation of symptom scales does not necessarily reflect good psychotherapy effects in patients' experiences. As such, the outcome classifications masked clinically relevant processes that can be highly relevant to instruct clinical practice on positive outcome and the role of psychotherapy (see Dattilio et al., 2010). Clearly, a more nuanced understanding of the relationship between therapy effects and patients' outcome or changes is required, calling for process-oriented research into the mechanisms of change (see Carey et al., 2006) that are often not limited to the influences of therapy alone (see De Smet et al., 2018; 2019).

Strengths, Limitations and Future Directions

The present study is the first large-scale mixed methods study focusing on both subgroups “recovery” and “improvement”, allowing for the contribution to the construction of more specified knowledge that is currently lacking in outcome research and the understanding of good outcome (in general, e.g., McElvaney & Timulak, 2013; and in the particular case of major depression, e.g., Rottenberg et al., 2018). The findings of this study can guide further usage and interpretations of clinically significant outcome in the treatment of depression. However, further testing of these findings in other samples and populations is needed to allow for more generalized conclusions. Moreover, further research focusing on investigating the differences and similarities between the various groups distinguished by the JT classification of clinical significance (i.e., recovery, improvement, no change and deterioration; Jacobson & Truax, 1991) would be highly instructive to fully grasp the value of this outcome classification in light of patients’ experiences (see De Smet et al., 2019 for non-improved patients’ experiences and De Smet, Cornelis et al., 2019 for deteriorated patients’ experiences).

In interpreting the findings of this study, it is important to address the implications of the research design (see Figure 1). While the first analyst was aware of patients’ outcome classification while performing the analyses, the second and third analyst were not during the initial phase of the analysis. Pre-existing knowledge inevitably impacts the analyses and therefore, masking patients’ outcome classifications, for part of the research team was expected to counter preliminary knowledge and to stay open for findings that contrast pre-existing assumptions. Starting from a qualitative research framework, it could be argued that patients’ outcome classifications provide essential contextual knowledge. This was, however, considered in the final stages of the analyses when all analysts were aware of outcome classifications. Furthermore, it can be assumed that differences between the recovered and improved groups would have been less apparent if interviews had been studied in random order and regardless of patients’ outcome, as pre-existing knowledge inevitably impacts the analyses (Stiles, 1993).

Adhering to a mixed methods philosophy and given our explicit aim to elucidate the meaning of the differentiated JT classifications, we however relied on the existing groups based on the outcome evaluation to scrutinize the experiences within these groups. An important complement to our findings and recommendation for further research would be to conduct the analyses starting from patients' experiences and examining whether the JT classification is in line with patients' perspectives in a second step. Generally, more fine-grained investigation of the relationship between measured outcome and patients' experiences is warranted and could benefit greatly from case comparison studies (Dattilio et al., 2010).

In general, the participants in our study form a representative outpatient population seeking psychotherapy in private practice for what can be considered one of the most prevalent mental health problems worldwide. However, our participants resembled a homogenous and local (predominately Caucasian, Flemish) group of patients. Specific (e.g., cultural, ethnic) or more diverse groups of patients should be the focus of complementing research. Despite the exclusion criteria guiding this study, which excluded acute suicidal patients and patients with primary substance abuse, comorbidity was allowed and as such patients in this study resembled a clinical reality (see Hirschfeld, 2001). However, bias in the selection of patients willing to participate in the study and socially desirable answers cannot be excluded completely. The therapy duration was fixed at 16-20 sessions and we evaluated patients' outcome shortly after therapy. For some patients this duration may have been insufficient to provide significant changes and outcome experiences may have been different if evaluated after a longer period. Further follow-up is required to evaluate patients' progress, especially in the case of depression, which is typically considered a recurring mental condition although little is known about those cases who show long-term improvement (Rottenberg et al., 2018).

This study explicitly aimed to address the much-used classification tool of clinical significance (Jacobson & Truax, 1991) and primary outcome method (De Los Reyes et al.,

2011) as dominantly applied in standard outcome research and RCTs in particular. The influence of the selected measure (i.e., BDI) on our findings is, however, inevitable. We can for instance expect that the BDI benefitted the outcome evaluation of patients in CBT over those in PDT given the more explicit focus on symptom reduction in the former treatment (see Levitt, Stanley, Frankel & Raina, 2005 for a critical discussion on this matter). As such, different patients might have been included as "good outcome cases" in the study if different measures (e.g., therapy-specific scales) were used. Replication with other measures and other selection criteria could thus be a valuable complement to our observations. Conclusions on the relation between quantitative and qualitative appraisals of outcome drawn from this study consequently also do not speak for the entire field of quantitative outcome evaluation that undoubtedly has evolved in the past decades (see for instance the development of personalized questionnaires, e.g., Elliott et al., 1999). Nonetheless, we can consider the choice for the BDI as a purposive case as it is one of the most frequently used measure of depression symptom severity. Moreover, our conclusions can inform on consequences related to pre-post standard outcome evaluation in general and the relevance of including underlying patients' experiences in general. Finally, although of indispensable importance, the chosen patient perspective can only shed light on outcome from one perspective. As previous research suggests different perspectives do not always accord (see Altimir et al., 2011), future research must aim at the incorporation of multiple perspectives (e.g., the therapist, researcher, society or family) in order to obtain a more complete picture of the clinical meaning of outcome.

Conclusion

The aim of this study was to contribute to the understanding of RCT outcome, by scrutinizing the experienced changes of patients that were marked as recovered and improved following psychotherapy. From patients' perspectives, good outcome can be understood as feeling empowered, finding personal balance and encountering ongoing struggle, indicating an ongoing

process and variation in experience. As such, the JT classification of "good outcome" could not account for some patients' contrasting (more pessimistic) outcome experiences, especially in the group of patients marked as "improved". Also, the JT indices did not allow us to fully grasp the multi-dimensional nature of outcome as experienced from patients' perspectives. Consequently, it is recommended that the statistical indications of clinical meaningfulness are interpreted warily and ideally contextualized within personal narratives. Further research on the phenomenon of change and good outcome is required and should aim at integrating multiple perspectives and methods in line with the multi-dimensional phenomenon under study.

References

Altimir, C., Krause, M., de la Parra, G., Dagnino, P., Tomicic, A., Valdés, N., … Vilches, O. (2010). Clients', therapists', and observers' agreement on the amount, temporal location, and content of psychotherapeutic change and its relation to outcome. *Psychotherapy Research, 20*(4), 472–487. https://doi.org/10.1080/10503301003705871

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Arlington, VA: American Psychiatric Association Publishing.

Atkins, D. C., Bedics, J. D., Mcglinchey, J. B., & Beauchaine, T. P. (2005). Assessing clinical significance: Does it matter which method we use? *Journal of Consulting and Clinical Psychology*, *73*(5), 982–989. https://doi.org/10.1037/0022-006X.73.5.982

Barrett, M. S., Chua, W.-J., Crits-Cristoph, P., Gibbons, M. B., Casiano, D., & Thompson, D. (2008). Early withdrawal from mental health treatment: implications for psychotherapy practice. *Psychotherapy (Chicago Ill), 45*(2), 247–267. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2762228/

Beck A.T., Steer R.A., & Brown G.K. (1996). *Manual for the Beck Depression Inventory–II.* San Antonio, TX: Psychological Corporation

Beckstead, D. J., Hatch, A. L., Lambert, M. J., Eggett, D. L., Goates, M. K., & Vermeersch, D. A. (2003). Clinical significance of the Outcome Questionnaire (OQ-45.2). *The Behavior Analyst Today, 4*(1), 86. https://doi.org/10.1037/h0100015

Binder, P.-E., Holgersen, H., & Nielsen, G. H. (2010). What is a "good outcome" in psychotherapy? A qualitative exploration of former patients' point of view. *Psychotherapy Research, 20*(3), 285–294. https://doi.org/10.1080/10503300903376338

Blanton, H., & Jaccard, J. (2006). Arbitrary metrics in psychology. *The American Psychologist*, *61*(1), 27–41. doi: 10.1037/0003-066X.61.1.62

Bockting, C., & Huibers, M. (2011). Protocollaire behandeling van patiënten met een depressieve stoornis; In G. Keijsers, A. van Minnen, & K. Hoogduin (Eds), *Protocollaire behandelingen voor volwassenen met psychiatrische klachten (Volume 1).* Amsterdam: Boom.

Bohart, C. A., & Wade, A. M. (2013). The client in psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behaviour change* (6th Ed.) (pp. 219-257). Hoboken, New Jersey: John Wiley & Sons, Inc.

Boyatzis, R. E. (1998). *Transforming qualitative information: thematic analysis and code development.* London, England: Sage

Braakmann, D. (2015). Historical paths in psychotherapy research. In O. C. G. Gelo, A. Pritz & B. Rieken (Eds.), *Psychotherapy research: foundations, process and outcome* (pp. 39-66). Vienna, Austria: Springer.

Carey, T. A., Carey, M., Mullan, R. J., Murray, L. K., & Spratt, M. B. (2006). Psychological Change: What changes and how does it occur? A critical review. *Counselling Psychology Review, 21*(4) 28-38

Carey, T. A., Carey, M., Stalker, K., Mullan, R. J., Murray, L. K., & Spratt, M. B. (2007). Psychological change from the inside looking out: A qualitative investigation. *Counselling and Psychotherapy Research*, *7*(3), 178–187. https://doi.org/10.1080/14733140701514613

Carey, T. A., & Stiles, W. B. (2016). Some Problems with Randomized Controlled Trials and Some Viable Alternatives. *Clinical Psychology & Psychotherapy*, *23*(1), 87–95.

Charmaz, K. (2014). *Constructing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage.

Cuijpers, P., Karyotaki, E., Weitz, E., Andersson, G., Hollon, S. D., & van Straten, A. (2014). The effects of psychotherapies for major depression in adults on remission, recovery and improvement: A meta-analysis. *Journal of Affective Disorders*, *159*, 118–126. https://doi.org/10.1016/j.jad.2014.02.026

Dattilio, F. M., Edwards, D. J. A., & Fishman, D. B. (2010). Case studies within a mixed methods paradigm: Toward a resolution of the alienation between researcher and practitioner in psychotherapy research. *Psychotherapy: Theory, Research, Practice, Training*, *47*(4), 427–441. https://doi.org/10.1037/a0021181

De Geest, R. M., & Meganck, R. (in press). How do time limits affect our psychotherapies? A literature review. *Psychologica Belgica.*

Delgado-Romero, E. A., & Romero-Shih, A. I. (2016). Patricia Arredondo: Creating a Pathway for Cultural Empowerment. *The Counseling Psychologist*, *44*(8), 1212–1235. https://doi.org/10.1177/0011000016683943

De Los Reyes, A., Kundey, S. M. A., & Wang, M. (2011). The end of the primary outcome measure: A research agenda for constructing its replacement. *Clinical Psychology Review*, *31*(5), 829–838. https://doi.org/10.1016/j.cpr.2011.03.011

De Smet, M. M., & Meganck, R. (2018). Understanding Long-term Outcome from the Patients' Perspective: A Mixed Methods Naturalistic Study on Inpatient Psychotherapy. *Psychologica Belgica, 58*(1), 276. doi: 10.5334/pb.432

De Smet, M. M., Meganck, R., Vannieuwenhove, K., Truijens, F. L., & Desmet, M. (2019). No Change? A Grounded Theory Analysis of Depressed Patients' Perspectives on Non-improvement in Psychotherapy. *Frontiers in Psychology, 10*(588). doi: 10.3389/fpsyg.2019.00588

De Smet, M.M., Cornelis, S., Acke, E., Reitske, M., Truijens, F., & Desmet, M. (2019). Making Sense of Deterioration in Psychotherapy: An Interpretative Phenomenological Analysis of Depressed Patients' Perspectives (manuscript forthcoming)

Ekroll, V. B., & Rønnestad, M. H. (2017). Processes and changes experienced by clients during and after naturalistic good-outcome therapies conducted by experienced psychotherapists. *Psychotherapy Research*, *27*(4), 450–468. https://doi.org/10.1080/10503307.2015.1119326

Elliott, R., Mack, C., & Shapiro, D. (1999). *Simplified Personal Questionnaire Procedure*. Toledo, OH: University of Toledo

Elliott, R., Slatick, E., & Urman, M. (2001). Qualitative change process research on psychotherapy: Alternative strategies. In J. Frommer & D. L. Rennie (Eds.), *Qualitative psychotherapy research: Methods and methodology* (pp. 69–111). Lengerich, Germany: Pabst Science.

Evans, I. M., & Fletcher, A. (2013). Focus on change: implications for the understanding and evaluation of psychological interventions. *Psychologia, 56*(2), 113–130. https://doi.org/10.2117/psysoc.2013.113

First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B.W. (2002). *Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition. (SCID-I/P).* New York: Biometrics Research, New York State Psychiatric Institute

Gianakis, M., & Carey, T. A. (2008). A review of the experience and explanation of psychological change. *Counselling Psychology Review, 23*(3), 27-38.

Giorgi, A., & Giorgi, B. M. (2003). The descriptive phenomenological psychological method. In P. M. Camic, J. E. Rhodes & L. Yardley (Eds.), *Qualitative research in psychology: expending perspectives in methods and design* (pp. 243-274). Washington, DC: American Psychological Association

Gladis, M. M., Gosch, E. A., Dishuk, N. M., & Crits-Christoph, P. (1999). Quality of life: expanding the scope of clinical significance. *Journal of Consulting and Clinical Psychology*, *67*(3), 320–331. http://dx.doi.org/10.1037/0022-006X.67.3.320

Glaser, B. G., & Strauss, A. L. (1987). *The discovery of grounded theory: Strategies for qualitative research*. Hoboken, New Jersey: Transaction publishers

Hesse-Biber, S. N. (2010). *Mixed methods research, merging theory with practice.* Guilford press, New York

Hill, C. (2004). *Helping skills: facilitating exploration, insight, and action* (2nd Ed.). Washington, DC: American Psychological Association.

Hill, C. E., Knox, S., Thompson, B. J., Williams, E. N., Hess, S. A., & Ladany, N. (2005). Consensual qualitative research: An update. *Journal of Counseling Psychology, 52,* 196–205. doi:10.1037/0022-0167.52.2.196

Hirschfeld, R. M. A. (2001). The Comorbidity of Major Depression and Anxiety Disorders: Recognition and Management in Primary Care. *Primary Care Companion to The Journal of Clinical Psychiatry*, *3*(6), 244–254.

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

Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy*, *15*(4), 336–352. https://doi.org/10.1016/S0005-7894(84)80002-7

Kazdin, A. E. (1977). Assessing the clinical or applied importance of behavior change through social validation. *Behavior modification*, *1*(4), 427-452.

Kazdin, A. E. (1999). The meanings and measurement of clinical significance. *Journal of Consulting and Clinical Psychology*, 332–339.

Kazdin, A. E. (2001). Bridging the Enormous Gaps of Theory With Therapy Research and Practice. *Journal of Clinical Child & Adolescent Psychology*, *30*(1), 59–66. https://doi.org/10.1207/S15374424JCCP3001_7

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

Knox, S., Schlosser, L. Z., Pruitt, N. T., & Hill, C. E., (2006). A qualitative examination of graduate advising relationships: The advisor perspective. *The Counseling Psychologist, 34,* 489–518. doi:10.1177/0011000006290249

Lambert, M.J. (2011). What have we learned about treatment failure in empirically supported treatments? Some suggestions for practice. *Cognitive and Behavioural Practice, 18*, 413–420. https://doi.org/10.1016/j.cbpra.2011.02.002

Lambert, M. J., Hansen, N. B., & Bauer, S. (2008). Assessing the Clinical Significance of Outcome Results. In A. M. Nezu & C. M. Nezu (Eds.). *Evidence-based outcome research. A Practical Guide to Conducting Randomized Controlled Trials for Psychosocial Interventions* (pp. 359-378). New York, NY: Oxford University Press

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. DOI:10.1080/10503300902849483

Leichsenring, F., & Schauenburg, H. (2014) Empirically supported methods of short-term psychodynamic therapy in depression – Towards an evidence-based unified protocol. *Journal of Affective Disorders, 169,* 128–143

Levitt, H. M., Stanley, C. M., Frankel, Z., & Raina, K. (2005). An Evaluation of Outcome Measures Used in Humanistic Psychotherapy Research: Using Thermometers to Weigh Oranges. *The Humanistic Psychologist*, *33*(2), 113–130. https://doi.org/10.1207/s15473333thp3302_3

Luborsky, L. (1984). *Principles of psychoanalytic psychotherapy. A manual for supportive expressive treatment*. NY: Basic Books.

McElvaney, J., & Timulak, L. (2013). Clients' experience of therapy and its outcomes in 'good' and 'poor' outcome psychological therapy in a primary care setting: An exploratory study. *Counselling and Psychotherapy Research*, *13*(4), 246–253. https://doi.org/10.1080/14733145.2012.761258

McLeod, J. (2013). Qualitative research. Methods and contributions. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed.) (pp. 49-84). New York, NY: John Wiley & Sons.

Meganck, R., et al. (2017). The Ghent Psychotherapy Study (GPS) on the differential efficacy of supportive-expressive and cognitive behavioral interventions in dependent and self-

critical depressive patients: study protocol for a randomized controlled trial. *Trials*, *18*(126), 1-11. DOI:10.1186/s13063-017-1867-x

Messer, S.B., & McWilliams, N. (2007). Insight in psychodynamic therapy: Theory and assessment, in L. G. Castonguay & C. E. Hill (Eds) *Insight in psychotherapy*. (pp. 9-30). Washington: American Psychological Association.

Midgley, N., Ansaldo, F., & Target, M. (2014). The meaningful assessment of therapy outcomes: Incorporating a qualitative study into a randomized controlled trial evaluating the treatment of adolescent depression. *Psychotherapy*, *51*(1), 128-137. Doi:10.1037/a0034179

Mohr, D. C., (1995). Negative outcome in psychotherapy: A critical review. *Clinical Psychology: Science and Practice, 2*(1), 1– 27. doi:10.1111/j.1468-2850.1995.tb00022.x

Mortelmans, D. (2013). *Handboek kwalitatieve onderzoeksmethoden (4th ed)*. Leuven, Belgium: Acco

Nilsson, T., Svensson, M., Sandell, R., & Clinton, D. (2007). Patients' experiences of change in cognitive-behavioural therapy and psychodynamic therapy: a qualitative comparative study. *Psychotherapy Research*, *20*, 37-41.

Ogles, B. M., Lunnen, K. M., & Bonesteel, K. (2001). Clinical significance: History, application, and current practice. *Clinical Psychology Review, 21*(3), 421-446. https://doi.org/10.1016/S0272-7358(99)00058-6

Parker, I. (2007). Critical psychology: What it is and what it is not. *Social and personality psychology compass, 1*(1), 1-15.

Ratcliffe, M. (2014). *Experiences of depression: A study in phenomenology*. OUP Oxford.

Rottenberg, J., Devendorf, A. R., Kashdan, T. B., & Disabato, D. J. (2018). The Curious Neglect of High Functioning After Psychopathology: The Case of Depression. *Perspectives on Psychological Science*, *13*(5), 549–566. https://doi.org/10.1177/1745691618769868

Slade, M., & Longden, E. (2015). Empirical evidence about recovery and mental health. *BMC Psychiatry*, *15*(1). https://doi.org/10.1186/s12888-015-0678-4

Stiles, W. B. (1993). Quality Control in Qualitative Research. *Clinical psychology Review, 13,* 593-618. doi: 10.1016/0272-7358(93)90048-Q

Timulak, L. (2007). Identifying core categories of client-identified impact of helpful events in psychotherapy: A qualitative meta-analysis. *Psychotherapy Research*, *17*, 310–320. https://doi.org/10.1080/10503300600608116

Timulak, L. (2010). Significant events in psychotherapy: An update of research findings. *Psychology and Psychotherapy: Theory, Research and Practice*, *83*(4), 421–447. https://doi.org/10.1348/147608310X499404

Tingey, R. C., Lambert, M. J., Burlingame, G. M., & Hansen, N. B. (1996). Assessing clinical significance: Proposed extensions to method. *Psychotherapy Research, 6*, 109–123.

van der Does, A.J.W. (2002). *BDI-II-NL. Handleiding. De Nederlandse versie van de Beck Depression Inventory-2nd edition.* Lisse: Harcourt Test Publishers.

Werbart, A., von Below, C., Brun, J., & Gunnarsdottir, H. (2015). "Spinning one's wheels": Nonimproved patients view their psychotherapy. *Psychotherapy Research*, *25*(5), 546-564. Doi: 10.1080/10503307.2014.989291

World Health Organization (2017). Depression and other common mental disorders: Global health estimates. Retrieved from http://apps.who.int/iris/bitstream/10665/254610/1/WHO-MSD-MER-2017.2

Zimmerman, M. A. (1995). Psychological empowerment: Issues and illustrations. *American Journal of Community Psychology*, *23*(5), 581–599.

Zimmerman, M., Martinez, J. A., Attiullah, N., Friedman, M., Toba, C., Boerescu, D. A., & Rahgeb, M. (2012). Why do some depressed outpatients who are in remission according to the Hamilton Depression Rating Scale not consider themselves to be in remission? *Journal of Clinical Psychiatry*, *73*(6), 790-795.

https://doi.org/10.1002/da.21987

IV. Change Processes underlying "Good Outcome": A Mixed Methods Study on Recovered and Improved Patients' Experiences in Psychotherapy for Major Depression

Abstract

Aim: Exploring change processes underlying "good outcome" in psychotherapy for major depression. We examined the perspectives of patients who "recovered" and "improved" (Jacobson & Truax,1991) over the course of CBT and PDT. Method: A mixed methods study was conducted on data from an RCT, comprising patients' pre-post outcome scores on the BDI-II and patients' post-therapy client change interviews. We selected the 28 patients who recovered, and 19 who improved in depression symptom severity. A grounded theory analysis was conducted and resulted in an integrative conceptual model. Results: According to recovered and improved patients' perspectives, change follows from an interaction between therapy, therapist, patient and extra-therapeutic context. Both helping and hindering influences were mentioned within all four influencing factors. Differences between recovered and improved patients point at the role of patients' agency and patients' internal and external obstacles. However, patients marked as "improved" described heterogeneous experiences. Conclusion: From patients' perspectives, various change processes underlie "good outcome" that do not necessarily imply an "all good process". This supports a holistic, multidimensional conceptualization of change processes in psychotherapy and calls for more fine-grained mixed-methods process-outcome research[23].

[23] De Smet, M. M., Meganck, R., Truijens, F., De Geest, R, Cornelis, S., Norman, U, & Desmet, M. (2019). Change Processes underlying "Good Outcome": A Mixed Methods Study on Recovered and Improved Patients' Experiences in Psychotherapy for Major Depression. *Psychotherapy Research* (manuscript accepted for publication pending revision)

Introduction

While an abundance of evidence supports the efficacy of psychotherapy, the processes underlying these beneficial results largely remain a black box (Cuijpers et al., 2014; Kazdin, 2009). Randomized Controlled Trials (RCTs) are challenged to increase their value for clinical practice (Westen, Novotny & Thompson-Brenner, 2004), as it is not only important *whether* treatment produces good outcome, but rather *how* positive changes come about (Greenberg, 1999; Kazdin, 2009). While outcome studies have predominantly been occupied with the pre-post, often symptom-based evaluation of patients' functioning, the focus is increasingly directed to an understanding of the underlying trajectory of change between these two points (Evans & Fletcher, 2013; Kazdin, 2009; McElvaney & Timulak, 2013).

The line of research concerned with these more complex process-outcome questions is known as "change process research" (CPR), which focuses on "identifying, describing, explaining, and predicting the effects of the processes that bring about therapeutic change'' (Greenberg, 1986, p. 4; see Elliott, 2010). RCTs especially, provide little information on the mechanisms through which treatments work, or the (specific) nature of the relationship between patients and therapy or therapists, CPR is considered an important complement to the "gold standard" in psychotherapy research (Dattilio, Edwards & Fishman, 2010; Elliott, 2010). With varying success, some features of the patient, the therapist and their interaction, as well as treatment activities predicting therapeutic outcome (e.g., the therapeutic relationship) have been identified (see Orlinsky, Rønnestad, & Willutzki, 2004). Nonetheless, a thorough evidence-based explanation of how or why interventions produce change (i.e., the mechanisms through which treatments operate) is lacking (Cuijpers, Reijnders, & Huibers, 2019; Kazdin, 2009). Research on the mechanisms of change is complicated by its inherent complexity. It is widely agreed that the idea of a single mechanism leading to a single outcome is too simplistic, as one influence can have multiple outcomes and similar outcomes can be reached via distinct pathways. In general, a linear understanding of change in psychotherapy – like the direct cause-

effect relations that are assumed in clinical trials (Carey et al., 2006; Greenberg, 1999) – is challenged (Kazdin, 2009; Tryon, 2005).

Elliott (2010) discusses four major approaches to study change processes that make it possible to understand these more complex interrelations. The "qualitative helpful factors design", or qualitative research in general, is in that regard believed to be an important means for in-depth examination of complex phenomena like change process and for the contextualization of research findings within, for instance, patients' lived experience (Levitt, Pomerville & Surace, 2016; McLeod, 2013). As patients can be considered "the site of change" (Greenberg, 1991, p.10) rather than the passive recipients of treatment, their perspective on the processes that lead to change is of indispensable value for both research and clinical practice, for example as a means of identifying helping and hindering factors (see Elliott, 2010; McLeod, 2013). An exceptional meta-synthesis by Levitt and colleagues (2016) on patients' experiences with psychotherapy found that "being known and cared for supports clients' ability to agentically recognize obstructive experiential patterns and address unmet vulnerable needs" (p.817). The findings point at the importance of the (professional role of the) therapist, the stimulation of patients' agency and adjustment of interventions to patients' needs as well as discussing differences between patient and therapist.

However, patients' voices are rarely taken into consideration when studying psychotherapy outcome, and consequently, this association remains largely uncomprehended. Moreover, the few mixed methods studies that have attempted to relate patients' experiences regarding therapy and change to measured outcome suggest that "good" and "poor" outcome are not as distinct as the categorization implies (e.g., McElvaney & Timulak, 2013). Patients in both groups mention similar experiences regarding outcome (e.g., ongoing struggles after treatment termination; see De Smet et al., 2019; in press), as well as psychotherapy experiences (see McElvaney & Timulak, 2013). Hence, in order to understand the processes associated with

the outcomes of psychotherapeutic treatment, more elaborative and mixed method research is needed**.**

Clinically, understanding the processes leading to beneficial therapy outcome is key to developing effective treatments and improving therapeutic care. However, while different mechanisms are assumed to cause "non-improvement" and "deterioration" in the case of "poor outcome" (Lambert, 2011), it remains unclear whether the same applies to "good outcome", typically understood as patients who "recover" and "improve" over the course of therapy (De Smet et al., in press). In a recent paper, Rottenberg and colleagues (2018) argued that more research on good outcome is especially warranted in the case of major depression, which is considered to be one of the most prevalent mental conditions worldwide (WHO, 2017). Limited attention has been allotted to the understanding of high functioning after treatment, in line with the dominant perception of depression as a *chronic* mental illness with poor prognosis. Consequently, knowledge lacks on how and why certain patients show sustainable successful outcome after experiencing depression. To address this neglect, Rottenberg and colleagues (2018) argue that research is required to pay more attention to the mechanisms leading to successful outcome and to expand the focus on symptomatic evaluation.

Therefore, in the present study we utilize the representative case of Major Depression to explore the change processes underlying good outcome as experienced by patients. In doing so, we rely on the widely used classification by Jacobson and Truax to identify good outcome and examine the perspectives of those patients who fall within the categories of "recovery" (i.e., reliable change and evolution to the non-clinical range) and "improvement" (i.e., reliable change but remaining in the clinical range) based on self-reported depression symptoms.[24] The

[24]The Jacobson-Truax classification (1991) is the most widely used definition of "clinical significance", which is a statistical indication of the clinical relevance of measured change in scores on an individual-patient level (i.e., whether a decrease in symptoms would be noticeable in actual clinical care). The JT-method relies on a twofold evaluation of outcome: Reliable change in pre-to-post treatment scores and evolution from dysfunctional to functional distribution (i.e., passing the clinical cut-off). Accordingly, patients are classified into four possible outcome categories: "Recovery" (reliable change and passing the clinical cut-off), "Improvement" (reliable change but remaining in the clinical range), "Non-improvement" (no reliable change)

study evolves around two central research questions: 1) what do patients, who show recovered and improved outcome, report as contributions to change? 2) what differences can be observed between experiences of patients classified as either recovered or improved? With the first research question, we aim to build onto the line of research looking into patients' experiences of psychotherapy (e.g., helping and hindering aspects; McLeod, 2013). Yet, as change processes are not necessarily exclusively happening within or due to the treatment (Carey et al., 2006; De Smet & Meganck, 2018), we expand the focus to include patients' experiences of what occurs outside of therapy as well. The second research question addresses the need for more specified mixed methods research on outcome in psychotherapy and particularly how it relates to patients' experiences (see De Smet et al., 2019; McElvaney & Timulak, 2013). Ultimately, we aim to provide insight into the change processes associated with successful (pre-to-post) RCT outcome findings from patients' perspectives. In doing so, this study build onto findings from the analyses of patients' perspectives on outcome and their experienced changes over the course of therapy, see De Smet, Meganck, De Geest et al. (2019; cf. method section).

Method

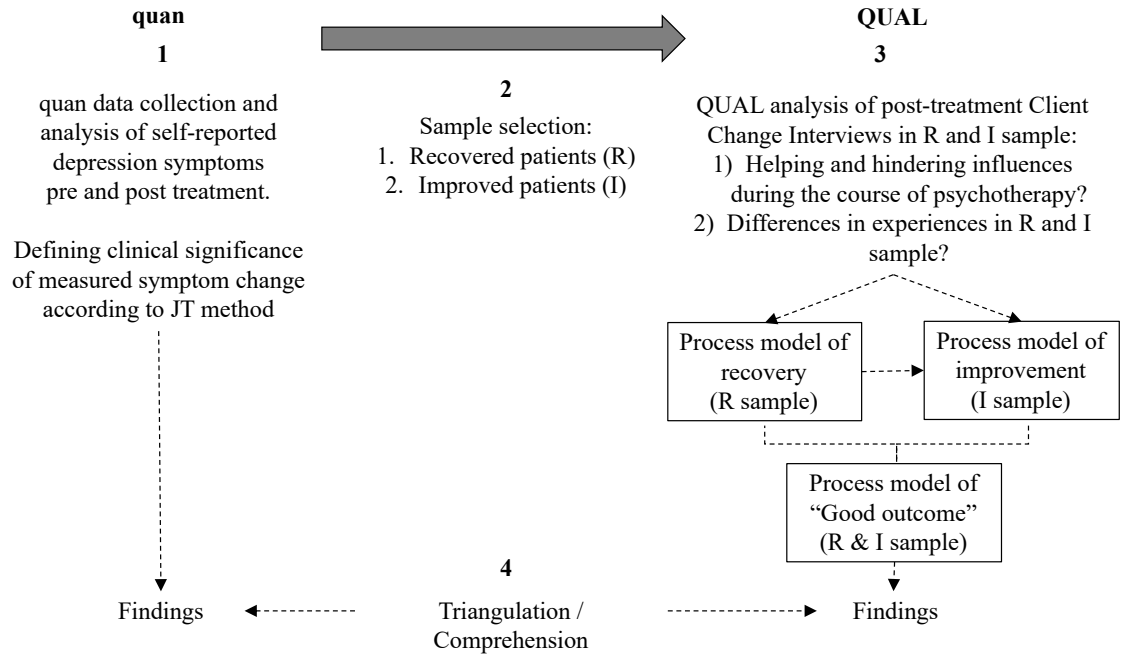
Design

An explanatory sequential mixed methods study (Hesse-Biber, 2010) was conducted, comprising a statistical test of clinical significance in measured symptom changes, as well as a qualitative analysis of patients' perspectives on change processes (Figure 1). In accordance with the mainstream primary outcome method (De Los Reyes, Kundey, & Wang, 2011), a psychometrically sound outcome measure was used to determine patients' outcome (see instruments). Both quantitative and qualitative data were gathered simultaneously; analyses were conducted independently and were integrated at the interpretation phase (Hesse-Biber,

or "Deterioration" (reliable change in the negative direction). "Recovery" is considered the gold standard for determining clinically significant change, although "improvement" is typically considered an indication of "good outcome" as well (see De Smet et al., 2019).

2010). Recovered and improved patients' interviews were studied separately and consecutively in order to investigate the statistical differentiation of good outcome. In a final stage, an integrative conceptual model was built, based on the analyses of both samples (see qualitative analysis below).

Figure 1. Explanatory Sequential Mixed Methods Design



Note. The design can be summarized as "quan → QUAL": The qualitative analyses follow the quantitative outcome evaluation yet become the most important focus of the explanatory study.

This study is part of a larger mixed methods research project on recovered and improved patients' experiences of outcome, psychotherapy and change processes. The study presented in this manuscript builds onto the inquiry of recovered and improved patients' perspectives on outcome and their experienced changes over the course of therapy. See De Smet, Meganck, De Geest et al. (2019) for an exploration of the experienced changes of recovered and improved patients. Both studies were conducted with the same samples of patients.

Setting

This study is situated within the Ghent Psychotherapy Study (GPS; Meganck et al., 2017), an RCT on the treatment of major depression. The main inclusion criterion was a diagnosis of Major Depressive Disorder (APA, 2013). Exclusion criteria comprised a primary diagnosis of substance abuse, acute psychosis and suicidal ideations. Patients were randomly assigned to time-limited Cognitive Behavioural Therapy (CBT) or Psychodynamic Therapy (PDT). Patients' progress was evaluated using questionnaires accompanying every therapy session, interviews were conducted prior to treatment, around the eighth therapy session and after treatment termination. The follow-up period of the study spans two years (ongoing), consisting of four interviews and quantitative assessments. This study was approved by the Ethical Committee of the University Hospital of Ghent University (Belgium; EC/2015/0085). All participants gave written informed consent prior to participating in the study.

Treatment

The investigated patients were allocated to either CBT and PDT for major depression, two types of therapy can be distinguished based on their directive (i.e., CBT) and explorative (i.e., PDT) style of interventions. Both treatments were manualized and consisted of 16-20 weekly sessions of 45 minutes. The CBT manual was based on the Cognitive-Behavioural Protocol for Depression (Bockting & Huibers, 2011). The PDT manual was based on the Supportive-Expressive Time Limited manual for Major Depressive Disorder by Luborsky (1984; Leichsenring & Schauenburg, 2014). Therapists had three to eight years of relevant clinical experience and training in CBT or PDT, and additionally received two days of training in the respective manuals, a patient to practice the treatment manual and research procedure under supervision, and bi-weekly supervision sessions.

Instruments

Beck Depression Inventory. The Beck Depression Inventory (BDI-II; Beck, Steer, & Brown, 1996; van der Does, 2002) is a measure of self-reported depression severity. The questionnaire consists of 21 items that are scored on a scale of 0 to 3 and is divided into a cognitive, somatic and affective subscale. A total score between 0-13 indicates minimal depression, 14-19 mild depression, 20-28 moderate depression, 29-63 severe depression. The Dutch version of the questionnaire, BDI-II-NL, shows good validity and reliability (Van der Does, 2002).

Semi-structured interview. An adjusted version of the semi-structured Client Change Interview (CCI; Elliott, Slatick & Urman, 2001) was administered to explore participants' experiences of therapy, the changes that occurred during therapy, and what they believe influenced these changes. Examples of some questions are: "What changes have you noticed since the start of therapy (e.g., in relation to others, at school/work, in your emotional wellbeing)?" and "How did therapy contribute to these changes?" All interviews were conducted by a research assistant at the university, approximately one week after therapy termination. Interviews were audiotaped for the purpose of verbatim transcription. Transcripts were analysed using Nvivo 11 and 12 (QSR International).

Quantitative Outcome Classification

Patients self-reported symptom severity was measured prior to therapy and one week after treatment ended and compared to Dutch norms (van der Does, 2002). The Jacobson and Truax (1991) method for outcome classification was used to classify participants in terms of improvement (i.e., reliable change) and recovery (i.e., clinically significant change). In order to reach reliable change for the BDI-II total score, a person must show a decrease in scores equal to or larger than 9.6. The cut-off between the clinical and nonclinical population for the Dutch BDI is set at 11.3 (based on the internal consistency of 0.92; van der Does, 2002). Thus, recovered patients showed a decrease in scores $\geq$ 9.6 and post-treatment score $<$ 11.3;

improvement implied a decrease in scores ≥ 9.6). Additionally, a post-hoc independent samples t-test was conducted to compare the mean depression pre-scores for the recovered and improved group (see participants).

Participants

From the one-hundred participants included in the RCT, seventy-four provided complete post data and were thus considered for the current sample selection. All patients showing recovery and improvement on the BDI-II (Van der Does, 2002) at treatment termination were selected (i.e., purposeful sample and total population sampling strategy). Patients with a pre-score ≤ 11.3 (i.e., clinical cut-off) and patients who dropped out (i.e., patient-initiated premature termination of therapy) within four sessions of treatment (in line with Barrett et al., 2006), were excluded (n=2). This resulted in a sample of 28 recovered participants (CBT: n=20; PDT: n=8) and 19 improved patients (CBT: n=7; PDT: n=12).

The recovered sample consisted of 18 women and ten men, ranging in age between 20 and 60 (M=33; SD=10.8). Out of the 28 patients, 17 were employed, four were unemployed and three patients had temporarily interrupted their work due to sick leave. The improved sample consisted of twelve women and seven men, ranging in age between 21 and 61 (M=40; SD=12.6). Out of the nineteen patients, twelve patients indicated they were employed, one patient was unemployed, and one had interrupted work due to sick leave. In both samples, previous treatment consisted of psychotherapy (for 22 recovered and 16 improved patients) and medication (for nine recovered and six improved patients). During the trial, the average number of therapy session was 16 (range 7-20) in the recovered sample and 18 (range 6-20) in the improved sample. All patients had a primary diagnosis of Major Depression Disorder at the start of treatment. Main secondary diagnoses were (as observed in the recovered and improved

sample respectively): Anxiety (*n*=6; *n*=5), Phobias (*n*=7; *n*=4), PTSD (*n*=0; *n*=4) and OCD (*n*=1; *n*=4)[25].

Table 1 summarizes the average score on the BDI-II (Beck, Steer, & Brown, 1996; van der Does, 2002) pre- and post-therapy in the recovered and improved sample.

Table 1. Pre-Post Outcome Scores on the BDI-II: Recovered and Improved Sample

	Recovered sample						**Improved sample**					
	Total (*n*=28)		CBT (*n*=20)		PDT (*n*=8)		Total (*n*=19)		CBT (*n*=7)		PDT (*n*=12)	
Total scores	M	SD (range)	M	SD (range)	M	SD (range)	M	SD (range)	M	SD (range)	M	SD (range)
Start therapy	29.2	7.4 (16-46)	29.7	7.2 (16-46)	28	7.37 (20-38)	38.5	7 (30-53)	36.3	4.6 (30-45)	39.8	7.6 (30-53)
End therapy	4.7	3 (0-9)	4.5	2.8 (0-9)	5.1	3.30 (0-9)	22.4	7.9 (13-41)	19.7	4.4 (13-22)	24	8.8 (14-36)

Note: Total scores on the BDI-II (van der Does, 2002): 0-13: minimal depression, 14-19: mild depression, 20-28: moderate depression, 29-63 severe depression. RCI: ≥ 9.6; clinical cut-off: < 11.3.

A statistically significant difference in scores for the recovered (M = 29.18, SD = 7.42) and improved (M = 38.52, SD = 7.14) group; t (45) = -4.30, p = .00 (two-tailed) at the start of treatment indicates that the improved sample scored significantly higher prior to therapy. The magnitude of the differences in the means (mean difference = -9.35, 95% *CI:* -13.72 to -4.97) was large (Cohen's d = 1.28). At the end of treatment, all recovered patients scored below the clinical cut-off (indicating minimal depression); the improved ranged from severely depressed to mildly depressed (resulting in a moderate depression mean score).

Qualitative analysis: Grounded Theory

[25]Other Axis 1 diagnoses present in one or two patients in the recovered group included: secondary substance abuse, eating disorder, dysthymia; and in the improved group: eating disorder, secondary substance abuse, somatic symptoms, pain disorder, hypochondria (according to SCID for DSM-IV-TR; First, Spitzer, Gibbon & Williams, 2002).

The explorative and interpretative method of grounded theory was conducted, as it allows for the construction of new rationales grounded in patients' interview data, which corresponds to our aim of understanding change processes (Charmaz, 2014; Glaser & Strauss, 1987). Analyses were conducted by the first author in dialogue with the fourth author; the second author functioned as auditor throughout the process (Hill et al., 2005; see below). The first author selected the samples and interviews corresponding to the sequential explanatory design outlined above (see Figure 1). The second and fourth author were aware of the focus of the study (i.e., good outcome cases and the existence of a recovered and improved sample of patients), but unaware of patients' specific outcome classification (i.e., whether patients belonged to the recovered or improved sample, nor raw BDI-II scores) during the open coding phase of the data analysis. This measure was taken to ensure that the researchers' perspectives remained open to findings contrasting potential pre-existing expectations. From the axial coding phase on (see below), all researchers were conscious of which themes belonged to the recovered or improved group. For a description of the analysis, see also De Smet, Meganck, De Geest et al. (2019).

Pertaining to the characteristics of grounded theory, several stages of analysis were completed in a cyclic manner before arriving at the final results (Charmaz, 2014). Analyses were first conducted on interviews of the recovered sample; interviews of the improved group were analysed subsequently. Categories were rephrased iteratively, or new ones were created to ensure that particular experiences in both samples were not overlooked. In the final stage of the analyses, central themes of both samples were defined to convey the experiences (see Figure 1). The *open coding phase*, consisted of extensively reading transcripts from patients' interviews to identify relevant parts (i.e., meaning units; Giorgi & Giorgi, 2003) relating to the research questions. This phase resulted in an initial list of codes, formulated with the intent to remain close to the narrative of patients. In the *axial coding phase*, these codes were thematically organized. In both samples, several "influencing factors" (hereafter referred to as

"factors") were extracted, referring to the role of therapy, therapist, patient, context and situational aspects. The *selective coding phase*, during which the theory was further refined by assembling the themes in four main factors (i.e., therapy, therapist, patient and extra-therapy context), resulted in a specification of the different influences, referred to as "change processes" observed within each factor. First, sample-specific subthemes were constructed aiming to grasp the experiences expressed in each sample (see Figure 1). In a next step, the resulting main and subthemes from both samples were strategically compared in order to formulate overarching factors and themes that could grasp the experiences of both the recovered and improved patients (i.e., underlying "good outcome"). Ultimately, a conceptual model (see Figure 2) was constructed, depicting the explicit purpose and findings of the study, i.e., elucidating the processes underlying "good outcome" within the samples of patients classified as "recovered" and "improved". The results were outlined using illustrative quotes (recovered patients were indicated by numbers 1 to 28 and improved patients were given the letters A to S to guarantee anonymity).The salience of the resulting factors and change processes (presented in Table 2) were indicated by means of the nomenclature proposed by Hill et al. (2005) and criteria for larger samples by Knox et al. (2006): General: ≥ 90% of the participants (25–28 and 17-19 recovered and improved patients respectively); Typical: ≥ 50% and < 90% (14-24 and 10-16 patients); Variant: ≥ 20% and < 50% (6-13 and 4–9 patients); Rare: < 20% (1-5 and 1-3 patients).

Credibility checks were held at several stages of data analysis. The research team conducting the analyses was specifically assembled to ensure a balance in terms of familiarity with the interviews, academic and clinical training background and career phase, enabling a horizontal and open dialogue (see Parker 2007). Influences were countered by remaining close to patients' own words (i.e., using their own descriptions) following a grounded theory bottom-up approach, rather than a theoretical interpretation of patients' narratives. An interest in

idiosyncratic experience and multi-facetted understanding of psychotherapy and change clearly led the study. Yet, by making this explicit, consequences of implicit guiding assumptions were precluded. During the study, preconceptions were left aside (i.e., "bracketing") as far as possible, by applying focused interviewing strategies in order to elucidate patients' experiences and using reflective diary and memo writing (e.g., in Nvivo) to reflect on own ideas during the analyses. The second author participated as independent researcher (i.e., not involved in conducting and analysing the interviews) and audited the process to actively test assumptions (e.g., perceived differences and similarities between the samples were challenged using critical questions). Generally, data analyses were guided by a contextual perspective, departing from the assumption that the broader social context influences how patients give meaning to their experiences (Boyatzis, 1998).

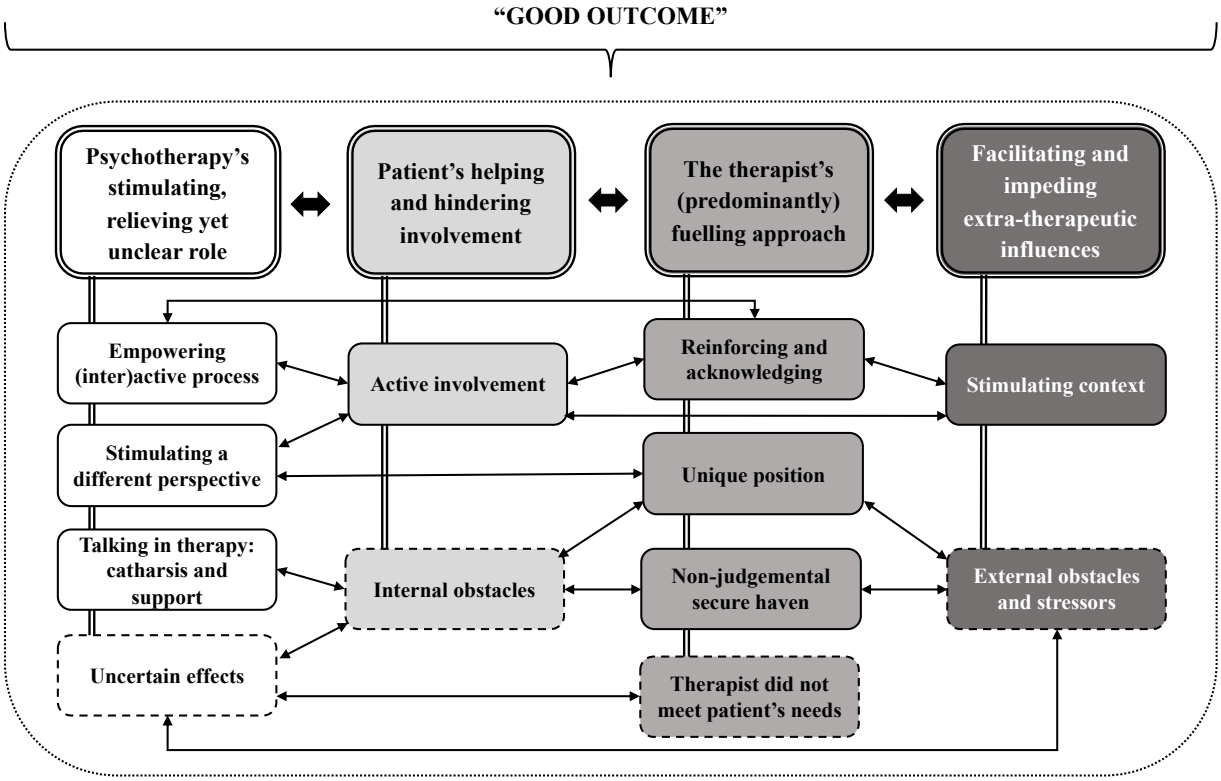
Results

The findings of the grounded theory analysis are visually presented in the conceptual model in Figure 2. Twelve operating processes resulted from the analyses and were subsequently clustered into four main influencing factors. Recovered and improved patients described: *I. Psychotherapy's stimulating, relieving yet unclear role, II. Patient's helping and hindering involvement, III. The therapist's predominantly fuelling approach,* and *IV. Facilitating and impeding extra-therapeutic influences*. These factors and underlying change processes must be understood as interconnected and non-mutually exclusive, meaning that all patients mentioned multiple processes operating at the same time. This implies that an interplay of processes facilitated change according to recovered and improved patients' perspectives, showing mutual influences rather than unidirectional causalities (see bidirectional arrows in Figure 2).

Table 2 presents the taxonomy of the main factors and change processes; they will be further described below. For both the recovered and the improved sample, the salience of the main factors and change processes is indicated with "general", "typical", "variant" or "rare"

depending on the number of cases to which they apply (based on Knox et al., 2006)[26]. Making it possible to address both important cross-sample similarities and within-sample particularities. We refer to recovered and improved samples across therapy conditions; when notable differences were observed for patients in CBT and PDT this is mentioned explicitly.

Figure 2. Conceptual Model of Change Processes Underlying "Good Outcome"



Note: The four main factors are depicted on the upside of the figure (i.e., rectangles surrounded by double lines). Change processes are depicted underneath the main factors and indicated by a corresponding tone of colour and connected with a double line. Dotted lines highlight the hindering influences within each factor. Bidirectional arrows show that mutual influences are assumed.

[26] General: ≥ 90% of the participants (25–28 and 17-19 recovered and improved sample respectively); Typical: ≥ 50% and < 90% (14-24 and 10-16 patients); Variant: ≥ 20% and < 50% (6-13 and 4–9 patients); Rare: < 20% (1-5 and 1-3 patients).

Table 2. Taxonomy of Change Processes Underlying "Good Outcome"

Change processes clustered into four overarching main factors	Recovered	Improved	*
I: Psychotherapy's stimulating, relieving yet unclear role	**General**	**General**	
Therapy stimulates a different perspective via multiple ways:	Typical	Typical	
Insights through talking, specific questions and reflection	*Typical*	*Typical*	
A confronting mirror	*Typical*	*Rare*	*
Making connections and overviews	*Variant*	*Rare*	*
Empowering (inter)active process	Typical	Variant	*
Talking leads to catharsis and gives support	Typical	Variant	*
Uncertain effects: insufficient, combined or unclear process	Typical	Typical	
II: Patients' helping and hindering involvement	**Typical**	**General**	*
Active involvement	Typical	Rare	*
Internal obstacles	Variant	Typical	*
III: The therapist's predominantly fuelling approach	**General**	**Typical**	*
Acknowledgement and reinforcement	Typical	Variant	*
Non-judgemental secure haven	Variant	Variant	
Unique position as person and professional	Rare	Variant	*
Therapist did not meet the patient's needs	Rare	Variant	*
IV: Facilitating and impeding extra-therapeutic influences	**Typical**	**General**	*
Stimulating context	Typical	Typical	
External obstacles or (situational) stressors	Variant	Typical	*

Note. *: highlights the differences between the recovered and improved sample in terms of the salience of the processes (according to Knox et al., 2006)

Main factor I: Psychotherapy's stimulating, relieving yet unclear role. Four therapy-related processes were described by recovered and improved patients, illustrating that in patients' experience, therapy on the one hand operated via tangible *stimulating* and *relieving* processes, yet on the other hand, the effects of therapy were sometimes also *unclear.* The different therapy processes are described in more detail hereafter.

Therapy stimulates a different perspective in multiple ways. Typically mentioned by both recovered and improved patients, therapy offered insights (e.g., regarding their relationships, past, problems) via the general process of *talking,* therapist's *(specific) questions* and *reflection*. It was important for patients to talk openly about their issues, to be asked stimulating questions about them by the therapist on which they then reflected on, in or outside of the therapy sessions. One patient describes gaining insight as follows:

> 'I think especially by actually talking about things, which I don't normally do. I typically form an idea in my head and keep on repeating it

> in my head, yet never say it out loud. And now, by doing precisely that, I start thinking about it in a different way.' (Patient C; PDT; Improved)

Two more specific processes were described, predominantly in the recovered sample and in rare cases by patients who improved. First, therapy typically presented recovered patients with a *confronting mirror,* generally mentioned in PDT and variantly in CBT. Patients stated that the therapist challenged them (e.g., certain preconceptions), which was deemed helpful to realize certain aspects of themselves that they had been unaware of. In that sense, the therapist and therapy were considered necessary to reveal one's blind spots[27].

> 'Another confrontation I received was like the sixth or seventh time [session] uhm, I sat there like very closed off, [answering] the therapist's questions. And she [said] I was not giving freedom to the process to find something else than the things I already knew.' (Patient 5; PDT; Recovered).

Moreover, therapy made patients see certain *connections* between previously diffused elements in their lives, and offered patients with *an overview,* typically for recovered patients, both in CBT and PDT. For instance, patients learned to see connections between their past and current problems (typically in PDT). Like one patient said: "The therapist basically helps you make the connections." Problems were visualized and outlined, and a more coherent story was built. In CBT this visualization was often made explicit via the use of schematic overviews.

An empowering (inter)active process. Therapy was valued for its active nature, as was typically mentioned by recovered patients, and variantly by patients in the improved sample. In CBT (generally) and PDT (typically), patients explained how they were actively involved in therapy (i.e., stimulated by the therapist) or it was considered empowering to take the lead in the therapy process. Especially patients in CBT (in both the recovered and improved sample) valued therapy for being *inter*active. The therapist participated actively and offered concrete tools and guidance.

[27] *Note:* "confronting mirror" is differentiated from the specific questions mentioned in "insights through talking, (specific) questions and reflection" given the explicit experience of being *confrontational*, which was not characteristic of the insight-stimulating questions.

> ‘In my previous treatment it was only action from my side, no response. (…) I need the interaction; therapists need to work *with* me.’ (Patient 7; CBT; Recovered)

Talking leads to catharsis and gives support. Attention was drawn to the importance of the simple act of talking during therapy, typically mentioned by patients who recovered and variantly by patients who improved. Recovered patients stated that, through talking in therapy, Pandora’s box was opened, emotions were set free and patients felt relieved. This effect of therapy was generally mentioned by patients in PDT and variantly by those in CBT.

> ‘A lot “came loose”. (…) I was completely stuck when I started therapy, I shed a lot of tears in therapy. (…) I’m less stuck now, things became more disentangled.’ (Patient 9; PDT; Recovered).

Apart from the relieving effect, improved patients also stated that talking in therapy gave them support (i.e., mentioned variantly), for instance, therapy was valued in terms of “finally” having *the opportunity and space* to talk freely (cf. unique position of the therapist): “It was important just to have someone to talk about it with every week, a sort of sounding board.” (Patient F; PDT; Improved).

Uncertain effects: insufficient, combined or unclear process. Despite the aforementioned positive effects of therapy, negative experiences were also voiced. In both samples, patients variantly expressed disappointment. For instance, CBT was experienced as too superficial (e.g., only focused on the present), PDT was seen as offering too little guidance (e.g., missing concrete tools) or overall, sessions were experienced as being too short, too repetitive or “just talking”. In the improved sample, patients variantly stated they were not sure how therapy had helped, describing a rather subtle process (e.g., “I feel therapy has helped, just no idea how”), most often in PDT (i.e., variantly versus rarely in CBT). Also, patients stated therapy had a rather variable impact, as formulated by one patient: “sometimes it’s about relevant things, sometimes it isn’t, and you can only gain from those relevant things.” (Patient I; PDT; Improved).

Both recovered (typically) and improved (variantly) patients indicated that therapy helped *in combination with other factors.* For instance, doing sports or yoga, taking time for oneself (e.g., during sick leave) and medication: "What lead to success for me personally, was a combination of things: psychotherapy, the psychiatrist and medication." (Patient 10; CBT; Recovered). Also, situational changes (e.g., finding a new partner; a new job; cf. extra-therapeutic influences) were mentioned as the main reasons for improvement, regardless of the effect of therapy: "It's not really therapy that helped me. (…) especially the situation I'm in improved." (Patient 20; CBT; Recovered).

Main factor II: Patients' helping and hindering involvement. Patients also reflected on their own role in the therapy and change process. This was characterized by *active involvement* (i.e., in the therapy and change process) and by hindering *internal obstacles.* As mentioned above, patients first of all felt actively involved in therapy (cf. (inter)active process). Second, and typically for recovered patients, an active position was reflected in the undertaking of other helping actions outside of therapy, for instance, doing sports, taking courses, deciding to take a break from work, which actively contributed to their own improvement (see combined effect of therapy, supra). This was only rarely the case for patients who were classified as "improved".

> 'The therapist did not give me assignments, but during the sessions, you start making your own tasks, like you arrive at the point: what am *I* going to do with this [i.e., what is discussed in therapy].' (Patient 26; CBT; Recovered)

However, patients were also hindered by *internal obstacles.* For instance, patients were convinced that it was impossible to change (e.g., the belief that problems are fundamental in nature), which was found variantly in recovered cases and typically for improved patients. Also, both the recovered and improved sample, variantly stated they had been unable to fully engage in therapy (e.g., difficulties with opening up), which was seen as hindering the therapy and

change process. For instance, one patient stated: “I don’t really have time to participate in something like this [therapy].”

Main factor III: The therapist’s predominantly fuelling approach. Although the role of the therapy and therapist are clearly intertwined, the person of the therapist was mentioned explicitly with regard to three benevolent processes. First, typically mentioned by recovered patients and variantly by patients who improved, the therapist was attributed the role of *acknowledging* one’s suffering and *reinforcing* one’s growth, which was experienced as empowering (together with therapy as an (inter)active process)

> ‘The therapist could tell me at times, “that’s not just nothing you’ve been through there, that’s a lot to bear” and yeah I thought indeed, perhaps, I never realized how heavy that was, and the therapist was able to put it into words.’ (Patient 27; CBT; Recovered)

Variantly in both samples, patients valued that the therapist *did not judge*, which allowed them to speak openly and show their true self. This *safe haven* in turn seemed to set the right condition for engaging in therapy (e.g., to be able to offer patients a confronting mirror).

> ‘[The therapist] could look at the situation in a neutral way, like telling me “it just happened, whether it was a good or bad thing, there are other lessons to be drawn from it than feeling guilty”. (…) I really needed to hear that.’ (Patient 21; CBT; Recovered).

Moreover, the therapist was described as being in a *unique position*, variantly mentioned by the improved patients, yet only rarely stated by recovered patients. For instance, the therapist was that one person who (finally) understood them, or the therapist as “a professional” was differentiated from normal social relationships. However, in correspondence to the unclear or sometimes disappointing effect of therapy, *therapists were not always able to meet patients’ needs* (rarely mentioned by recovered patients; variantly the case for improved patients). For instance, one patient said the therapist came across as inexperienced: “It was like she did not always know what to do with me.” Others did not fully understand a therapist’s opinion, had

difficulty with the therapist's approach (e.g., abrupt ending of the sessions when "the time was up"), or generally expected more from their therapist.

Main factor IV: Facilitating and impeding extra-therapeutic influences. Finally, patients in both samples talked about influences outside of the therapy room, again characterized by positive and negative aspects. *External obstacles or stressful circumstances* were seen as hindering to therapy, or as influencing patients' current mood (i.e., post-therapy) - typically mentioned by patients in the improved sample (typically in PDT and variantly in CBT). For instance, arguments with family members, employment uncertainty, chronic illness, or divorce, were ongoing during treatment and complicated the therapy and recovery process. Similar difficulties were variantly mentioned by recovered patients. However, social circumstances were also experienced as facilitating. Recovered and improved patients typically mentioned the importance of being *supported* by significant others during the therapy process and difficult times (e.g., having someone to lean on).

> 'If didn't have my boyfriend, I would not have started this [treatment] and then I still would have been feeling bad.' (Patient D; PDT; Improved).

Also, patients felt *empowered* by the positive responses they received on their changes (e.g., "my colleagues dare to trust me again, they feel I've calmed down"; Patient H; PDT; Improved). Moreover, patients' professional life took a central place in these extra-therapeutic experiences (both for the recovered and improved sample). For instance, work (e.g., having a job again) added meaning to patients' lives: "I have a goal again, a valuable way to spend my time." (Patient 4; PDT; Recovered). In rare cases, a lack of support (or even negative comments on being in therapy and becoming "different") caused patients unhelpful distress (see External obstacles).

In sum, mutual influences are assumed between the four main factors that together influenced change according to recovered and improved patients (Figure 2). Moreover, multiple processes within each factor were described, showing that both positive and negative influences

underlie “good outcome” in psychotherapy for major depression. The implications of these findings will be discussed hereafter.

Discussion

This study explored change processes associated with “good outcome”, based on experiences of patients whose outcome indicated recovery and improvement after CBT and PDT for major depression (Kazdin, 2009). In doing so, we expanded the understanding of the processes underlying successful (pre-to-post) RCT outcome findings (Elliott, 2010). This study builds onto the analyses of patients’ perspectives on outcome and experienced change over the course of therapy and we refer the reader to De Smet, Meganck, De Geest et al. (2019) for a comprehensive understanding of the recovered and improved patients’ perspectives on “good outcome”.

In this study, patients first and foremost mentioned a broad range of processes and changes, in line with previous qualitative research findings (McLeod, 2013). Moreover, besides the influence of therapy, patients acknowledged their own role, the therapist, and the influence of external obstacles and personal context outside of therapy (in line with De Smet & Meganck, 2018; De Smet et al., 2019). As such, our findings show that the observation of good pre-to-post psychotherapy outcome yields an interplay of multiple change processes, which supports a holistic and multidimensional understanding of change in psychotherapy (Levitt et al., 2016; McLeod, 2013). Interestingly, in light of the fact that we studied patients showing good psychotherapy outcome, also negative experiences were mentioned. Thus, from a patient perspective, “good outcome” did not necessarily correspond to an “all good process”.

Specifically, building onto the line of research concerned with helping processes and significant events in psychotherapy (see McLeod, 2013), patients in this study highlighted three central therapy-related processes. First of all, therapy stimulated a different perspective or insight, for instance, by means of specific questions, confrontation, seeing connections (e.g.,

between present and past events) and offering an overview. This corresponds to "pattern identification", the importance of "analysis of thoughts and assumptions" and "reflexivity" as found by Levitt et al. (2016). Second, therapy was experienced as an active and interactive process, which reflects the more practical component of therapy or "collaborative process" (Levitt et all., 2016) and the importance of empowerment (Timulak, 2007). Finally, therapy was experienced as relieving, especially through the cathartic act of talking (cf. Timulak, 2010).

Interestingly, patients in this study explicitly referred to their therapist as a unique influencing factor besides – although naturally related to – the role of therapy. This shows the importance of the therapist as the one setting the right condition for therapy to be helpful. For instance, therapists were experienced as reinforcing and acknowledging, and as such contributed to the empowering role of therapy. A "non-judgemental safe haven" or unique position on the other hand seemingly allowed patients to engage in the more difficult therapeutic work (e.g., confrontational insights; opening Pandora's box). Generally, the therapeutic relationship and the understanding and accepting role of the therapist are two important common factors in psychotherapy (e.g., trust, empathy; see Cuijpers et al., 2019) that were represented across both the recovered and improved sample in this study. A minority of the patients, however, felt their needs were not heard by the therapist, suggesting the occurrence of a rupture in the alliance (see Safran, Muran, & Eubanks-Carter, 2011).

When comparing recovered and improved patients' experiences, cross-sample resemblances suggest that similarities in experiences may outweigh the differences, as previously observed by McElvaney and Timulak (2013) when comparing good and poor outcome cases. However, notable differences between the samples point at the particular influence of patients' active involvement as well as internal and external obstacles. Regarding the former, especially recovered patients took on a more active role in their therapy and change process, similar to an "agentic activity" that is considered one of the most important patient

characteristics related to outcome (cf. Bohart & Wade, 2013, p. 245). As noted by McElvaney and Timulak (2013) we can wonder whether the position of the patients was facilitated through therapy or by the therapist (e.g., recovered patients mentioned more often that the therapist encouraged them; see Williams & Levitt, 2007 for agency facilitating principles), or whether recovered patients were already more empowered and ready to change before the start of therapy, in line with stages of change models. For instance, Norcross, Krebs and Prochaska (2011) differentiate an action stage "in which individuals modify their behaviour, experiences, and/or environment to overcome their problems" from contemplation in which patients are "still struggling to initiate change" (p.144). Internal and external obstacles (e.g., life circumstances) were more often mentioned by improved patients as interfering with the therapy process, potentially withholding patients from engaging in and benefitting fully from therapy (see Kazdin, Holland, Crowley, & Brenton, 1997).

Second, patients in the recovered sample provided more concrete processes with regard to the helping role of therapy (e.g., confronting mirror, connections and overview), which relates to the ability of "storying" psychotherapy experiences that has been suggested as important for sustaining positive therapy effects (see Adler, Wagner & McAdams, 2007, on the role of narratives, p. 1195). Patients in the improved sample, in contrast, described therapy in more general terms as being enriching, with more subtle effects, uncertain or even unclear contributions. That the majority of patients in the improved sample had engaged in PDT while recovered patients predominantly engaged in CBT, may offer an explanation here. As observed by Nilsson et al. (2007), treatment results were most clear to patients in CBT while the effects of PDT were less easy to grasp and more ambivalent, although both samples showed improvement. This points at differences in experiences depending on the directive (i.e., CBT) and explorative nature (i.e., PDT) of the treatments. Notably, however, patients who engaged in PDT and were classified as recovered provided clearer influences than did the improved

patients, again pointing at the importance of the patient's active involvement for understanding treatment effects.

Although this was not the main focus of our study, some observations regarding differences in the experiences of patients from CBT and PDT are interesting in light of previous findings, for instance by Nilsson et al. (2007). For both recovered and improved patients in CBT, the interactive nature of therapy was valued and experienced as empowering, in line with what Nilsson et al. (2007) described as the focused and active description of a directive therapy/therapist. Both CBT and PDT stimulated patients' insight, via schematic overviews in CBT and by connecting parts of patient's story into a more coherent narrative in PDT (i.e., a focus on 'self-understanding'; Messer & McWilliams, 2007), and by confronting patients with aspects of themselves by critical questions, both in CBT and PDT. The latter was especially found in PDT in the study by Nilsson and colleagues. The catharsis effect of talking, as typically mentioned in PDT, might resonate with the emotional processing in psychotherapy (Watson & Bedard, 2006), which is considered a crucial link between gaining insight and obtaining a positive outcome (Messer & McWilliams, 2007). While this was not explicitly described by Nilsson et al. (2007), they found patients valued PDT for "getting to the roots of things" in a working-through fashion.

Noteworthy, in line with the unclear (or sometimes even disappointing) effects of psychotherapy found in this study, the examination of recovered and improved patients' *outcome experiences* indicated that not all patients felt improved, and ongoing struggles were generally expressed in both the improved and recovered sample (see De Smet, Meganck, De Geest et al., 2019 on what "good outcome" means to patients). Moreover, especially the improved sample showed to be characterized by heterogeneity in experiences (ranging from feeling improved to not doing well at all), which questioned the classification of "good outcome" for a part of the patients. Hence, when interpreting the findings of the present study

on the general differences observed between recovered and improved patients, this heterogeneity should thus be taken into account (note for instance that the salience of the processes in the improved sample were often mentioned "variantly", rather than typical or general, illustrating the variety in experiences across the sample).

Our findings raise several important questions regarding the relationship between outcome (based on standardized, symptom-based outcome measures pre-post therapy) and patients' experiences of how changes came about. First, the observed differences within the samples in this study challenge the idea that all patients can be considered as moving towards a similar outcome following a similar (linear) trajectory of change (Kazdin, 2009). For instance, differences in patients' position in therapy or readiness for change prior to treatment (Norcross et al., 2011), life circumstances (Kazdin et al., 1997) and symptom severity (cf. the improved sample scored significantly higher prior to therapy; Clarkin & Levy, 2004) challenge the comparability of patients between and within samples. We can indeed wonder whether patients suffering from ongoing difficulties like chronic illness or divorce should aim towards the same outcome and require the same therapeutic processes as patients with different personal and contextual demands. Typically, this type of contextualizing information is omitted in RCTs in which randomization and standardization are assumed to rule out any confounding influences, however in doing so, the clinically meaningful context is also overlooked (Evans & Fletcher, 2013; Truijens et al., 2019).

Second, both samples, but especially the improved patients, challenged the idea that "good outcome" reflects "good therapy". For several patients, therapy was not considered helpful, or only in combination with other means. Noteworthy, moreover, is that these findings resemble the experiences of a *non*-improved sample of depressed patients (see De Smet, Meganck et al., 2019). Similar experiences in therapy were mentioned by patients who did not improve, as for instance therapy provided insights but also hit a certain limit (e.g., inability to

change). Also, similar to improved patients, internal and external obstacles were mentioned as important hindering influences. Hence, comparison of these patients' experiences suggests that symptomatic change in itself may not be a sufficient criterion to differentiate samples of patients when aiming to elucidate the processes that lead to meaningful change. Indeed, symptom reduction has shown to be only minimally meaningful to patients themselves (see De Smet et al., in press). Conversely, patients who deteriorate in symptom severity do not necessarily consider therapy as failing, nor is symptom increase straightforwardly experienced as a negative evolution (see De Smet, Cornelis et al., 2019).

Hence, if we wish to unravel which processes lead to successful outcome and how the role of therapy, therapist, patient and external context could be best organized for optimal treatment results, it is of essence to study the relationship between processes and outcome in more detail. This goes back to the recommended refinement of research questions suggested by Kiesler (e.g., "which therapist behaviours are more effective with which types of patients?"; 1966, p.113) who challenged the uniformity assumptions in psychotherapy research. In the specification of such moderators and mediators (see Kazdin, 2009), systematic case studies offer an important contribution, as they allow for the idiosyncratic contextualization of outcome and change processes (Datillio et al., 2010). However, a crucial – but often overlooked – link in this endeavour is the conceptualization of outcome or meaningful change in psychotherapy itself (see Carey et al., 2006). As argued by Evans and Fletcher (2013), "only if you know what outcome you want, you can evaluate whether the treatment has achieved it or not" (p.117). Hence, if the patient can be considered "the site of change" (Greenberg, 1991; p.10), a multidimensional study and holistic approach to process-outcome research seems warranted, in accordance to patients' holistic experience of change (cf. Levitt et al., 2016) and psychotherapy (cf. present study).

Strengths, Limitations and Future Directions

This study does not allow or aim to draw causal conclusions regarding the processes that lead to good outcome. Rather, it builds on the tradition of change process research that aims to bridge the gap between outcome research on the one hand and process research on the other (Elliott, 2010). We particularly address the highly influential framework of the JT classification of clinical significance as a means to offer a patient-centred indication of clinical meaningfulness. In that sense, the present study is the first large-scale mixed methods study focusing on both subgroups of "good outcome" (i.e., "recovery" and "improvement") to explore the underlying change processes. This contributes to the construction of more specified knowledge that is currently lacking in understanding the mechanisms in significant outcome (in general, McElvaney & Timulak, 2013; and in the particular case of major depression, Rottenberg et al., 2018). The resulting tentative conceptual model offers suggestions for further research and contributes to the conceptualization or theory-building regarding the important change processes in psychotherapy. In this regard, our findings build onto important meta-analytic knowledge (see Levitt et al., 2016; Timulak, 2007; 2010) yet expands its focus by offering a contextualized perspective on influences that expand the therapeutic context. Further research should focus on investigating the various outcome categories distinguished by the JT classification of clinical significance (i.e., recovery, improvement, no change and deterioration; Jacobson & Truax, 1991) in order to scrutinize the relationship between standard statistical outcome evaluation and underlying processes as experienced by patients in more detail. Systematic comparative case studies could be of great value here (Dattilio et al., 2010).

In the present study, therapy was provided in context of a standardized RCT setting, which inevitably influenced the findings in comparison to a naturalistic setting (Westen, et al., 2004). Nonetheless, the participants in our study correspond to a representative outpatient population seeking psychotherapy in private practice for what can be considered one of the most prevalent mental health problems worldwide. Specific (e.g., cultural, ethnic) or more

diverse groups of patients could be the focus of complementing research. Similarly, given the exclusion criteria of the study, we cannot draw conclusions regarding more severe and acute psychopathologies, although a significant number of patients in our study showed comorbid disorders, which resembles clinical practice (cf. Hirschfeld, 2001). Interestingly, our findings show striking resemblances to what was mentioned by patients showing successful outcome in a naturalistic study (Ekroll & Rønnestad, 2017).

The present study departed from one outcome measure to define patients pre-post outcome in accordance to common practices in standard outcome research (see De Los Reyes et al., 2011). The influence of the selected measure (i.e., BDI) on our findings is inevitable, as we can for instance expect that this symptom-based measure benefitted the outcome evaluation of patients in CBT over those in PDT given the more explicit focus on symptom reduction in the former treatment (see Levitt, Stanley, Frankel & Raina, 2005). A slightly different constellation of "good outcome cases" could have been included in the study if different measures were used. Replication with other measures and other selection criteria could thus be a valuable complement to our observations, especially if we wish to elucidate therapy-specific processes and their relation to outcome (cf. Levitt et al., 2016). Conclusions on the relation between quantitative measures and patients' (qualitative) experiences also do not speak for the entire field of quantitative outcome evaluation that undoubtedly has broadened its focus (beyond symptom-evaluation) in the past decades (see for instance the development of personalized questionnaires, e.g., Elliott et al., 1999). Nonetheless, we can consider the choice for the BDI as purposive in this case, as it is one of the most frequently used measure of depression symptom severity in outcome studies. Furthermore, our conclusions can inform on the consequences related to pre-post standard outcome evaluation in general and the relevance of including underlying patients' experiences. Finally, complementing the perspectives of patients in this study, future research must aim at the incorporation of multiple perspectives

(e.g., the therapist, researcher, society or family; Altimir et al., 2010) in order to obtain a complete understanding of the processes operating in and outside the context of psychotherapy.

Conclusion

From patients' perspectives, "good outcome" does not correspond to an "all good process". Recovered and improved patients described various therapy, therapist, patient and external contextual processes as operating at the same time, moreover, characterized by both helping and hindering influences. This supports a holistic, multidimensional conceptualization of change processes in psychotherapy and calls for more fine-grained mixed-methods research on the relationship between change processes and outcome.

References

Adler, J. M., Wagner, J. W., & McAdams, D. P. (2007). Personality and the coherence of psychotherapy narratives. *Journal of Research in Personality*, *41*(6), 1179–1198.

Altimir, C., Krause, M., de la Parra, G., Dagnino, P., Tomicic, A., Valdés, N., … Vilches, O. (2010). Clients', therapists', and observers' agreement on the amount, temporal location, and content of psychotherapeutic change and its relation to outcome. *Psychotherapy Research, 20*(4), 472–487.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Arlington, VA: American Psychiatric Association Publishing.

Barrett, M. S., Chua, W.-J., Crits-Cristoph, P., Gibbons, M. B., Casiano, D., & Thompson, D. (2008). Early withdrawal from mental health treatment: implications for psychotherapy practice. *Psychotherapy (Chicago Ill), 45*(2), 247–267.

Beck A.T., Steer R.A., & Brown G.K. (1996). *Manual for the Beck Depression Inventory–II.* San Antonio, TX: Psychological Corporation

Bockting, C., & Huibers, M. (2011). Protocollaire behandeling van patiënten met een depressieve stoornis; In G. Keijsers, A. van Minnen, & K. Hoogduin (Eds), *Protocollaire behandelingen voor volwassenen met psychiatrische klachten (Volume 1).* Amsterdam: Boom.

Bohart, C. A., & Wade, A. M. (2013). The client in psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behaviour change* (6th Ed.) (pp. 219-257). Hoboken, New Jersey: John Wiley & Sons, Inc.

Boyatzis, R. E. (1998). *Transforming qualitative information: thematic analysis and code development.* London, England: Sage

Carey, T. A., Carey, M., Mullan, R. J., Murray, L. K., & Spratt, M. B. (2006). Psychological change: what changes and how does it occur? A critical review. *Psychological Change*, 9.

Charmaz, K. (2014). *Constructing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage.

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

Cuijpers, P., Karyotaki, E., Weitz, E., Andersson, G., Hollon, S. D., & van Straten, A. (2014). The effects of psychotherapies for major depression in adults on remission, recovery and improvement: A meta-analysis. *Journal of Affective Disorders*, *159*, 118–126.

Cuijpers, P., Reijnders, M., & Huibers, M. J. H. (2019). The Role of Common Factors in Psychotherapy Outcomes. *Annual Review of Clinical Psychology, 15,* 207-231

Dattilio, F. M., Edwards, D. J. A., & Fishman, D. B. (2010). Case studies within a mixed methods paradigm: Toward a resolution of the alienation between researcher and practitioner in psychotherapy research. *Psychotherapy: Theory, Research, Practice, Training*, *47*(4), 427–441.

De Los Reyes, A., Kundey, S. M. A., & Wang, M. (2011). The end of the primary outcome measure: A research agenda for constructing its replacement. *Clinical Psychology Review*, *31*(5), 829–838.

Denzin, N. K., & Lincoln, Y. S. (2005). *The sage handbook of qualitative research.* Thousand Oaks, CA: Sage

De Smet, M., & Meganck, R. (2018). Understanding Long-term Outcome from the Patients' Perspective: A Mixed Methods Naturalistic Study on Inpatient Psychotherapy. *Psychologica Belgica, 58*(1), 276. DOI: https://doi.org/10.5334/pb.432

De Smet, M., Meganck, R., Vannieuwenhove, K., Truijens, F., & Desmet, M. (2019). No Change? A Grounded Theory Analysis of Depressed Patients' Perspectives on Non-improvement in Psychotherapy. *Frontiers in Psychology, 10(588). doi.10.3389/Ffpsyg.2019.00588*

De Smet, M. M., Meganck, R., De Geest, R., Norman, U. A., Truijens, F., & Desmet, M. (in press). What "good outcome" means to patients: Understanding Recovery and Improvement in Psychotherapy for Major Depression from a Mixed-Methods Perspective. *Journal of Counseling Psychology*. Http://dx.doi.org/10.1037/cou0000362

De Smet, M. M., Cornelis, S., Acke, M., Meganck, R., Truijens, F., & Desmet, M. (2019). Making sense of deterioration in psychotherapy: An interpretative phenomenological analysis of depressed patients' perspectives. *Manuscript submitted for publication.*

Ekroll, V. B., & Rønnestad, M. H. (2017). Processes and changes experienced by clients during and after naturalistic good-outcome therapies conducted by experienced psychotherapists. Psychotherapy Research, 27(4), 450–468.

Elliott, R. (2010) Psychotherapy change process research: Realizing the promise. *Psychotherapy Research, 20*(2), 123-135, DOI: 10.1080/10503300903470743

Elliott, R., Mack, C., & Shapiro, D. (1999). *Simplified Personal Questionnaire Procedure*. Toledo, OH: University of Toledo

Elliott, R., Slatick, E., & Urman, M. (2001). Qualitative change process research on psychotherapy: Alternative strategies. In J. Frommer & D. L. Rennie (Eds.), *Qualitative psychotherapy research: Methods and methodology* (pp. 69–111). Lengerich, Germany: Pabst Science.

Evans, I. M., & Fletcher, A. (2013). Focus on change: implication for understanding and evaluation of psychological interventions. *Psychologia, 56*(2), 113–130.

First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B.W. (2002). *Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition. (SCID-I/P).* New York: Biometrics Research, New York State Psychiatric Institute

Giorgi, A., & Giorgi, B. M. (2003). The descriptive phenomenological psychological method. In P. M. Camic, J. E. Rhodes & L. Yardley (Eds.), *Qualitative research in psychology: expending perspectives in methods and design* (pp. 243-274). Washington, DC: American Psychological Association

Glaser, B. G., & Strauss, A. L. (1987). *The discovery of grounded theory: Strategies for qualitative research*. Hoboken, New Jersey: Transaction publishers

Greenberg, L. S. (1986). Change process research. *Journal of Consulting and Clinical Psychology, 54*, 4-9.

Greenberg, L. S. (1991). Research on the process of change. *Psychotherapy Research*, *1*(1), 3–16.

Greenberg, L. S. (1999). Ideal psychotherapy research: A study of significant change processes. *Journal of Clinical Psychology*, *55*(12), 1467–1480.

Hill, C. E., Knox, S., Thompson, B. J., Williams, E. N., Hess, S. A., & Ladany, N. (2005). Consensual qualitative research: An update. *Journal of Counseling Psychology, 52,* 196–205.

Hirschfeld, R. M. A. (2001). The Comorbidity of Major Depression and Anxiety Disorders: Recognition and Management in Primary Care. *Primary Care Companion to The Journal of Clinical Psychiatry*, *3*(6), 244–254.

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

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

Kazdin, A. E., Holland, L., Crowley, M., & Brenton, S. (1997). Barriers to treatment participation scale: Evaluation and validation in the context of child outpatient treatment. *Journal of Child Psychology and Psychiatry, 38,* 1051-1062.

Kiesler, D. J. (1966). Some myths of psychotherapy research and the search for a paradigm. *Psychological Bulletin, 65*(2), 110–136.

Knox, S., Schlosser, L. Z., Pruitt, N. T., & Hill, C. E., (2006). A qualitative examination of graduate advising relationships: The advisor perspective. *The Counseling Psychologist, 34,* 489–518.

Lambert, M.J. (2011). What have we learned about treatment failure in empirically supported treatments? Some suggestions for practice. *Cognitive and Behavioural Practice, 18*, 413–420.

Leichsenring, F., & Schauenburg, H. (2014) Empirically supported methods of short-term psychodynamic therapy in depression – Towards an evidence-based unified protocol. *Journal of Affective Disorders, 169,* 128–143

Levitt, H. M., Stanley, C. M., Frankel, Z., & Raina, K. (2005). An Evaluation of Outcome Measures Used in Humanistic Psychotherapy Research: Using Thermometers to Weigh Oranges. The Humanistic Psychologist, 33(2), 113–130.

Levitt, H. M., Pomerville, A., & Surace, F. I. (2016). A qualitative meta-analysis examining clients' experiences of psychotherapy: A new agenda. *Psychological Bulletin*, *142*(8), 801–830.

Luborsky, L. (1984). *Principles of psychoanalytic psychotherapy. A manual for supportive expressive treatment*. NY: Basic Books.

McElvaney, J., & Timulak, L. (2013). Clients' experience of therapy and its outcomes in 'good' and 'poor' outcome psychological therapy in a primary care setting: An exploratory study. *Counselling and Psychotherapy Research*, *13*(4), 246–253.

McLeod, J. (2013). Qualitative research. Methods and contributions. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed.) (pp. 49-84). New York, NY: John Wiley & Sons.

Meganck, R., et al. (2017). The Ghent Psychotherapy Study (GPS) on the differential efficacy of supportive-expressive and cognitive behavioral interventions in dependent and self-critical depressive patients: study protocol for a randomized controlled trial. *Trials*, *18*(126), 1-11.

Messer, S. B., & McWilliams, N. (2007). Insight in psychodynamic therapy: Theory and assessment. In L. G. Castonguay & C. Hill (Eds.), *Insight in psychotherapy* (pp. 9–29). Washington, DC: American Psychological Association.

Nilsson, T., Svensson, M., Sandell, R., & Clinton, D. (2007). Patients' experiences of change in cognitive-behavioural therapy and psychodynamic therapy: a qualitative comparative study. *Psychotherapy Research*, *20*, 37-41.

Norcross, J. C., Krebs, P. M., & Prochaska, J. O. (2011). Stages of change. In J. C. Norcross (Ed.), *Psychotherapy relationships that work: Evidence-based responsiveness* (2nd ed., pp. 279–300). New York, NY: Oxford University Press.

Orlinsky, D. E., Rønnestad, M. H., & Willutzki, U. (2004). Process and outcome in psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 307-389). New York: Wiley.

Rottenberg, J., Devendorf, A. R., Kashdan, T. B., & Disabato, D. J. (2018). The Curious Neglect of High Functioning After Psychopathology: The Case of Depression. *Perspectives on Psychological Science*, *13*(5), 549–566.

Safran, J. D., Muran, J. C., & Eubanks-Carter, C. (2011). Repairing alliance ruptures. *Psychotherapy, 48*(1), 80–87.

Stiles, W. B. (1993). Quality Control in Qualitative Research. *Clinical psychology Review, 13,* 593-618. doi: 10.1016/0272-7358(93)90048-Q

Timulak, L. (2007). Identifying core categories of client-identified impact of helpful events in psychotherapy: A qualitative meta-analysis. *Psychotherapy Research*, *17*, 310–320.

Timulak, L. (2010). Significant events in psychotherapy: An update of research findings. *Psychology and Psychotherapy: Theory, Research and Practice*, *83*(4), 421–447.

Truijens, F. L., Desmet, M., De Coster, E., Uyttenhove, H., & Deeren, B. (in press). When quantitative measures become qualitative story books. A phenomenological case analysis of the validity and performativity of questionnaire administration in psychotherapy research. *Journal of Qualitative Research for Psychology.*

Tryon, W. W. (2005). Possible mechanisms for why desensitization and exposure therapy work. *Clinical Psychology Review, 25,* 67- 95.

van der Does, A.J.W. (2002). *BDI-II-NL. Handleiding. De Nederlandse versie van de Beck Depression Inventory-2nd edition.* Lisse: Harcourt Test Publishers.

Watson, J. C., & Bedard, D. L. (2006). Clients' emotional processing in psychotherapy: A comparison between cognitive-behavioral and process-experiential therapies. *Journal of Consulting and Clinical Psychology*, *74*(1), 152–159.

Westen, D., Novotny, C. M., & Thompson-Brenner, H. (2004). The Empirical Status of Empirically Supported Psychotherapies: Assumptions, Findings, and Reporting in Controlled Clinical Trials. *Psychological Bulletin*, *130*(4), 631–663.

Williams, D. C., & Levitt, H. M. (2007). Principles for facilitating agency in psychotherapy. *Psychotherapy Research*, *17*(1), 66–82.

V. No Change? A Grounded Theory Analysis of Depressed Patients' Perspectives on Non-improvement in Psychotherapy[28]

Abstract

Understanding the effects of psychotherapy is a crucial concern for both research and clinical practice, especially when outcome tends to be negative. Yet, while outcome is predominantly evaluated by means of quantitative pre-post outcome questionnaires, it remains unclear what this actually means for patients in their daily lives. To explore this meaning, it is imperative to combine treatment evaluation with quantitative and qualitative outcome measures. This study investigates the phenomenon of non-improvement in psychotherapy, by complementing quantitative pre-post outcome scores that indicate no reliable change in depression symptoms with a qualitative inquiry of patients' perspectives. The study took place in the context of a Randomized Controlled Trial evaluating time-limited psychodynamic and cognitive behavioural therapy for major depression. A mixed methods study was conducted including patients' pre-post outcome scores on the BDI-II-NL and post treatment Client Change Interviews. Nineteen patients whose data showed no reliable change in depression symptoms were selected. A grounded theory analysis was conducted on the transcripts of patients' interviews. From the patients' perspective, non-improvement can be understood as being stuck between knowing versus doing, resulting in a stalemate. Positive changes (mental stability, personal strength and insight) were stimulated by therapy offering moments of self-reflection and guidance, the benevolent therapist approach and the context as important motivations. Remaining issues (ambition to change but inability to do so) were attributed to the therapy hitting its limits, patients' resistance and impossibility and the context as a source of distress. "No change" in outcome scores therefore seems to involve a "partial change" when considering the patients' perspectives. The study shows the value of integrating qualitative first-person analyses into standard quantitative outcome evaluation and particularly for understanding the phenomenon of non-improvement. It argues for more multi-method and multi-perspective research to gain a better understanding of (negative) outcome and treatment effects. Implications for both research and practice are discussed.

[28] De Smet, M. M., Meganck, R., Van Nieuwenhove, K., Truijens, F. L., & Desmet, M. (2019). No Change? A Grounded Theory Analysis of Depressed Patients' Perspectives on Non-improvement in Psychotherapy. *Frontiers in Psychology, 10.* https://doi.org/10.3389/fpsyg.2019.00588

Introduction

Negative outcome or nonresponse to treatment is undeniably part of clinical practice. It is estimated that 5 to 10% of patients deteriorates in therapy (Cooper, 2008; Lambert, 2013), and a proportion of 35 to 40% of the participants in clinical trials do not improve (Lambert, 2007). A better understanding of negative outcome and treatment effects is crucial for both research and clinical practice, yet outcome research has focused predominantly on capturing positive change and "what works", while less is known about non-improvement or what it actually means when treatments fail (Barlow, 2010).

There is no uniform understanding of negative outcome, nor is there agreement on the definition of treatment failure (Lambert, 2011; Lampropoulos, 2011). "Negative outcome" and "negative therapeutic effects" are often used as synonyms, although they do not have a one-on-one relationship, as negative outcome is not necessarily caused by therapy (Mays & Franks; 1985; Mohr, 1995). Depending on the perspective (e.g., patient, therapist, researcher), the type of outcome (e.g., symptoms, quality of life), measurement method (e.g., quantitative or qualitative) and time point (e.g., post treatment or follow-up) being used for treatment evaluation, the conception of outcome and treatment effects varies (Lampropoulos, 2011).

In outcome research, outcome and treatment effects are typically evaluated using statistical tests of significance that provide an indication of the reliability of the measured change. Statistical significance shows that an outcome difference is larger than could have been expected by mere chance. Clinical significance shows whether such a statistical effect is also clinically meaningful (i.e., change towards a normal level of functioning) (Jacobson et al., 1999; Lambert, Hansen, & Bauer, 2008; Lambert & Ogles, 2009; Ogles, Lunnen, & Bonesteel, 2001). Based on the Jacobson and Truax widely used method for clinical significance, outcome can be classified into four categories: 1) recovery (i.e., clinically significant change), 2) improvement (i.e., reliable change), 3) no reliable change and 4) deterioration (i.e., reliable change in the negative direction). Generally, the first category 'recovery' is taken as the gold standard

outcome and treatment goal: a reliable decrease in symptoms[29] and return to a non-clinical level of functioning. When neither criterion is met, it is concluded that patients remained "unchanged" in comparison to their level of functioning prior to treatment (Jacobson & Truax, 1991).

Despite the added value of clinical significance testing of measured changes, this type of statistical outcome classification cannot overcome the limitations that are voices for standard outcome research (Hill, Chui, Baumann, 2013). Quantitative pre-post outcome evaluation is criticized for relying predominantly on one-dimensional rating scales, most often symptom-based (Braakmann, 2015), and consequently, for offering only an incomplete approximation of the multi-dimensional nature of human functioning (Hill et al., 2013; Kazdin, 2001). The possible discrepancy between what is measured with outcome questionnaires and what is meaningful in patients' daily life has been problematized: a patient's outcome score might fall within the non-clinical range while it does not reflect the person's functioning (Kazdin, 1999; 2001). Real-life contextualization is necessary in order to make sense of what changes in scores (or the lack thereof) actually mean for an individual (Blanton & Jaccard, 2006; Kazdin, 2006). The latter is typically missing in large sample standardized outcome studies, and consequently, the dissemination of research findings into clinical practice generally fails (Kazdin, 2008).

The past decades have seen an accumulation of qualitative studies attempting to contribute to overcoming this research-practice barrier, gradually offering a more central role to the voice of patients (Levitt, Pomerville, & Surace, 2016). Qualitative research focussing on patients' experiences of outcome has provided a diverse picture of treatment-related changes (McLeod, 2011). Apart from symptomatic changes, alterations on the level of patients' self, life, interpersonal relations and self-understanding have been observed (e.g., Binder, Holgersen

[29] Note of nuance: as outcome is predominantly evaluated by means of symptom-based scales (Braakmann, 2015), throughout this paper we will refer to a decrease in symptoms, even so we acknowledge outcome measurement is not limited to symptom-based scales but can also contain measures of general wellbeing, satisfaction and interpersonal functioning.

& Nielsen, 2009). The largest strand of qualitative psychotherapy research has focused on patients' experiences of therapy, aiming to identify helping and hindering aspects (McLeod, 2013). Hindering elements in therapy that have been mentioned by patients are contra-productive therapist features (e.g., being unsure, absent or non-responsive, lack of direction and advise in therapy), patients' own difficulties to express or get in touch with their feelings and lack of commitment and motivation, and a lack of trust between patient and therapist (Paulson, Everall, & Stuart, 2001; von Below & Werbart, 2012) and so forth. On the other hand, a joint exploration of difficulties and experiencing warmth, understanding and empathy in the relationship with the therapist were found to be helpful for patients (Bohart & Wade, 2013; Lilliengren & Werbart, 2005; Timulak & Lietaer, 2001).

Interestingly, findings from qualitative outcome studies shine a somewhat more pessimistic light on psychotherapy outcome than is typically observed in quantitative studies. In general, patients tend to be more critical about therapy during interviews, for instance, expressing disappointment about unaltered core problems or ambivalence about the gains of therapy (McLeod, 2013). Moreover, research findings suggest that, patients' treatment satisfaction does not correspond to changes in outcome scores. Werbart et al. (2015), for instance, observed that only three out of twenty patients with a nonimproved or deteriorated outcome also clearly indicated to be dissatisfied about treatment.

Nonetheless, the association between quantitative and qualitative evaluations of therapy and outcome remains unclear (Timulak & Creaner, 2010). Mixed-methods studies have amassed in the past couple of years, though whether and how patients' experiences correspond to quantitative outcome evaluation is underexplored (McLeod, 2013). The few studies that have been executed differ in the extent to which qualitative and quantitative findings show an accord (see Svanborg, Bäärnhielm, & Wistedt, 2008 versus Klein & Elliott, 2006). The study of McElvaney and Timulak (2013) found only little differences between patients classified as

"recovered/ improved" and "unchanged/ deteriorated" regarding their experience of therapy. As the strict demarcation of "poor" and "good" outcome does not appear in qualitative inquiry, questions can be raised about how representative such a statistical distinction is for the clinical meaning of outcome for individual patients (see Lambert & Ogles, 2009).

So far, the meaning of negative or poor outcome – distinguished by means of standard outcome measures - in relation to patients' subjective experiences, remains underexplored. As non-improvement and worsening are likely distinct phenomena with potential different clinical implications (Lambert, 2011; Mohr, 1995), more focused investigations are required in order to grasp the phenomenon of non-improvement (in contrast to the approach of McElvaney & Timulak, 2013, who studied unchanged and deteriorated cases together). In past endeavours, most of the studies have focused on deterioration or other "extreme cases", yet little attention has been allotted to understanding treatment nonresponse or patient non-improvement specifically (Lambert, 2011). Given the observation that lack of improvement occurs in a significant number of cases, and considerably more frequently than deterioration, this lack of attention is striking (Lambert, 2007; 2013). Importantly, gaining a better understanding of cases who seemingly have not moved forward or backward, will contribute to a more thorough and nuanced understanding of treatment-response and outcome in general. More specifically, this nuanced understanding is pivotal to elaborate the clinical meaning of outcome for patients themselves.

The integration of multiple methods and specifically the comparison of quantitative and qualitative methods is an indispensable development for the field of psychotherapy research (Bowie, McLeod, & McLeod, 2011; McLeod, 2013). The current study therefore provides a mixed-method analysis of patients suffering from major depression. Major depression is one of the most prevalent mental disorders worldwide (WHO, 2017), and previous research has shown symptomatic evaluation of change alone cannot live up to the task of representing depressed

patients' experience of outcome (Zimmerman et al., 2006; 2012). Based on this representative case, the present study aims to complement quantitative pre-post outcome scores indicating no reliable change in depression symptoms with a qualitative inquiry of depressed patients' perspective. In doing so, we move beyond the level of description (i.e., a lack of change in symptom scores) and towards a level of in-depth understanding (i.e., patients' subjective experience). Finally, instead of adopting a single focus on experiences of outcome or experiences of therapy, the present study aims to understand their interrelation as well as the broader context of potential influences, as these are typically not limited to therapeutic features alone (De Smet & Meganck, 2018; Drisko, 2004).

The current study investigates how non-improvement in pre-to-post symptom severity can be understood in relation the experience of depressed patients themselves. We examine: 1) which potential changes patients have experienced, and which factors can help to explain these changes from their perspective; 2) which potential issues remained, and which factors can help explain these remaining issues according to patients; 3) how patients' perspective on non-improvement relates to the quantitative outcome evaluation of non-improvement (or no reliable change) in symptom severity. For the purpose of the study, the term "non-improvement" is used to indicate a specific definition of negative or poor outcome in accordance to the widely used statistical concept of a lack of reliable change in outcome scores (cf. Jacobson & Truax, 1991). We use this categorization as a starting point to be able to *broaden* this influential framework of understanding, by nuancing it based on patients' perspectives.

Method

An explanatory sequential mixed-methods study was conducted, comprising a quantitative pre-post outcome evaluation as well as a qualitative analysis of nonimproved patients' perspective. The study is 'explanatory' as the focus is on understanding non-improvement in-depth, and 'sequential' because, even though quantitative and qualitative data were gathered

simultaneously, both strands were analysed independently and integrated at the phase of interpretation. The design can be summarized as "quan → QUAL": The qualitative analyses build on the quantitative outcome evaluation yet becoming the most important focus of the explanatory study; 'the quantitative study (quan) is in service of the more dominant qualitative (QUAL) one' (Hesse-Biber, 2010, p.71). In the current study, a first phase comprised a quantitative outcome evaluation, based on which the target sample was selected. In a second phase, the corresponding interviews were qualitatively analysed. Integration and comparison of the two strands allowed for a better understanding of both the quantitative and qualitative outcome findings. Given the aim for in-depth exploration of patients' experienced changes, as well as understanding of the processes and factors that may explain those experienced changes, a grounded theory approach was selected as method of choice for the qualitative analyses (Strauss & Corbin, 1990). Grounded theory can be used to provide description and interpretation, with the aim to generate conceptual models that can consecutively be translated into further hypotheses (Charmaz, 2014; Fassinger, 2005). For our purposes, thus, this method seemed well-suited to build a thorough understanding of negative outcome and non-improvement from patients' perspective.

Setting

This study is based on data from the Ghent Psychotherapy Study (GPS), an RCT on the treatment of major depression; the trial has been registered on Open Science Framework (ISRCTN 17130982). For a specific description of the GPS context and methodology, we refer to the pre-registered study protocol (Meganck et al., 2017). Patients in this study were recruited via social media and general practitioners in the area of Ghent, Belgium. Patients included in the study qualified for a diagnosis of Major Depressive Disorder, measured by the Rating Scale for Depression (Hamilton, 1967) and Structured Clinical Interview for DSM-IV-TR (First, Spitzer, Gibbon, & Williams, 2002), both well-established and frequently used interview-based

instruments in depression studies (Nezu, Ronan, Meadows & McLure, 2000). The assessment interviews were conducted by six postgraduate research assistants trained in the respective procedures. Further eligibility criteria were sufficient knowledge of the Dutch language and age between 18 and 65; patients with a primary diagnosis of substance abuse, acute psychosis and suicidal ideations were excluded. Patients were randomly assigned to time-limited Cognitive Behavioural Therapy (CBT) or Psychodynamic Therapy (PDT). Patients progress was evaluated using questionnaires accompanying every session, interviews were conducted prior to treatment, around the eighth session and after treatment termination. The follow-up period of the study spans two years (ongoing) and consists of four interviews and quantitative assessment. This study was approved by the the Ethical Committee of the University Hospital of Ghent University (Belgium; EC/2015/0085). All participants gave written informed consent in accordance with the Declaration of Helsinki.

Treatment

Treatment consisted of CBT and PDT for major depression, two types of therapy that can be distinguished based on their directive (i.e., CBT) and exploration (i.e., PDT) style of interventions. Therapy was provided by one of four therapists in each approach. Both treatments were manualized and time-limited, consisting of 16-20 sessions. Treatment was delivered with an average frequency of one session per week; sessions lasted approximately 45 minutes. The CBT manual was based on the Cognitive-Behavioural Protocol for Depression by Bockting and Huibers (2011). The PDT manual was based on the Supportive-Expressive Time Limited manual for Major Depressive Disorder by Luborsky (1984; Leichsenring & Schauenburg, 2014). Therapists had an average age of 33 (SD=9.6) and had three to eight years of relevant clinical experience and training in CBT or PDT. In the study, all therapists received two days of training, one patient to practice the treatment manual and the research procedure under supervision, and bi-weekly supervision sessions throughout the study.

Instruments

Beck Depression Inventory. The Beck Depression Inventory (BDI-II-NL; Beck, Steer, & Brown, 1996; van der Does, 2002)30 is a measure of self-reported depression severity. The questionnaire consists of 21 items that are scored on a scale of 0 to 3 and is divided into a cognitive, somatic and affective subscale. A total score between 0-13 indicates minimal depression, 14-19 mild depression, 20-28 moderate depression, 29-63 severe depression. The questionnaire shows good validity and reliability (van der Does, 2002).

Semi-structured interview. An adjusted version of the semi-structured Client Change Interview (CCI; Elliott, Slatick & Urman, 2001) was administered. The interview guide was constructed to evoke participants' experiences of therapy, the changes they believe occurred during therapy, and what they believe influenced these changes, for instance, helping and hindering aspects of therapy. Every interview started with the open questions: "How are you doing in general?" and "How are you feeling compared to when you started therapy?" Subsequently, patients were asked more specifically about experienced changes: "Which changes have you noticed since the start of therapy (e.g., in relation to others, at school/work, in your emotional wellbeing)?" and the role of therapy or other factors: "How did therapy contribute to these changes?" and "What other factors (outside of therapy) do you think have contributed to these changes?" Patients were also explicitly asked about negative changes or lack of change: "Is there something that did not change or that you would like to change in the future?"; "Did something change in a negative sense during therapy?" All interviews were conducted at the psychology department (Ghent University, Belgium) in the week following therapy termination. Interviews lasted 60 minutes on average. Interviews were audiotaped, and transcripts were analysed using Nvivo 11 (QSR International).

[30] Given the focus of the current study on patients being treated for major depression, the BDI-II-NL was selected as the outcome measure in this study. A complete overview of all measures used in the GSP can be found in the study protocol (Meganck et al., 2017).

Quantitative Outcome Classification on the BDI-II-NL

Participants were classified in terms of reliable change and clinically significant change based on the Jacobson and Truax (1991) method for outcome classification. Patients self-reported symptom severity was measured prior to therapy and one week after treatment ended. The outcome scores of the patient population were compared to Dutch norms (van der Does, 2002). In order to reach reliable change for the BDI-II-NL total score, a person must show a decrease in scores equal to or larger than 9.6. The cut-off between the clinical and nonclinical population for the Dutch BDI is set at 11.3 (based on the internal consistency of 0.92; van der Does, 2002). This leads to four possible outcomes: Clinically significant change (CS; a decrease in scores equal to or larger than 9.6 and post-treatment score below 11.3), reliable change (RC; a decrease in scores equal to or larger than 9.6), no RC (a decrease or increase in scores < 9.6) and deterioration (an increase in scores equal to or larger than 9.6). In the total sample of the RCT (*n*=94), 31.9% (*n*=30) of the patients changed clinically significant, 20.2% (*n*=19) changed reliably, 23.4% (*n*=22) remained unchanged and 3.2% (n=3) deteriorated in scores on the BDI-II-NL; 21.3% (*n*=20) had missing outcome data (see Figure 1).

Participants

For the current study, patients showing no reliable change in pre-to-post outcome scores on the BDI-II-NL (van der Does, 2002) were included. We did not incorporate deteriorated patients based on the assumption that non-improvement and worsening are distinct phenomena with potential different clinical implications (cf. supra; Lambert, 2011; Mohr, 1995). For the same reason, we excluded patients who ended treatment prematurely (i.e., drop-out from treatment), which was defined as the patient-initiated premature termination of therapy within four sessions of treatment (in line with other commonly used definitions of drop-out; cf. Barrett et al., 2006; Wierzbicki & Pekarik, 1993) This resulted in the selection of 19 participants. The flowchart in Figure 1 gives an overview of the selection process for this study. The sample consisted of

twelve women and seven men ranging in age from 21 to 59 (*M*=34; *SD*=10.7). All patients were born in Belgium except for one patient who was born in the Netherlands; one patient had a parent of foreign origin. Table 1 gives an overview of the demographic information per patient. During the study, eight patients received CBT; eleven patients received PDT. Patients' average treatment duration was 17 sessions (range 6-20 sessions). All patients were diagnosed with major depression prior to treatment (comorbid Axis I diagnoses as assessed using the SCID for DSM-IV-TR are presented in Table 1).

Figure 1. Flow chart of sample selection

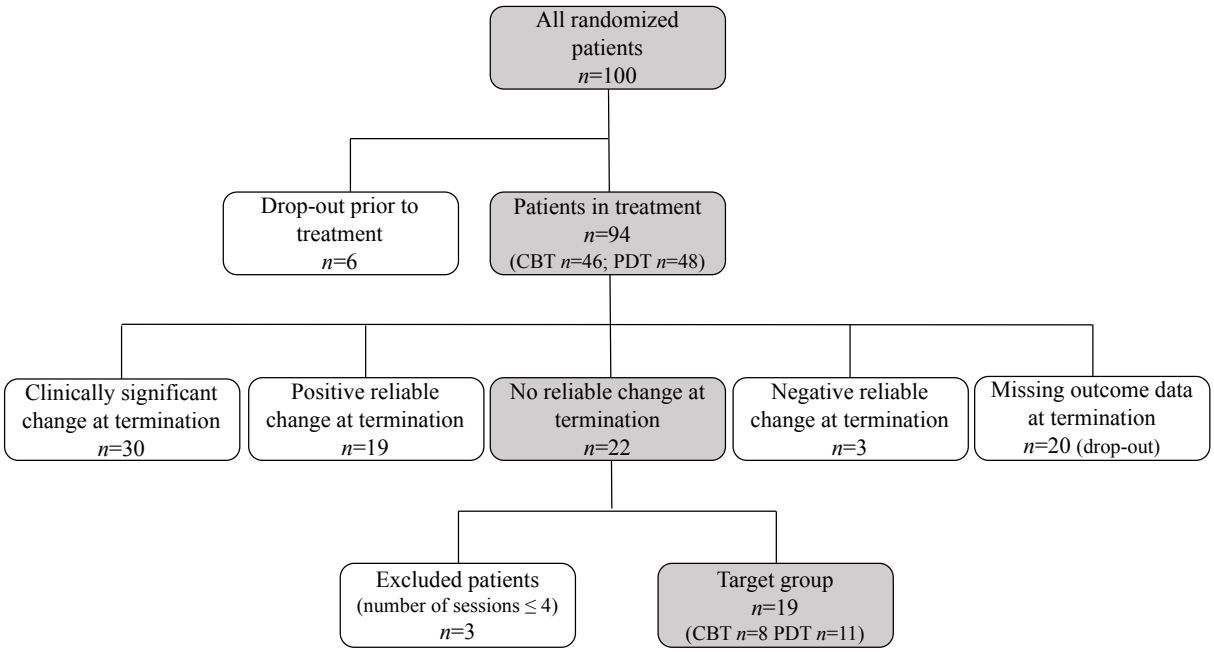


Table 1. Demographical information of patients in the sample

Pt	M/F	Age range	Marital status	Education level	Employment status	Therapy (n sessions)	Prior care	Comorbid diagnoses
A	F	35-40	Divorced	Higher	Employed	PDT (20)	Therapy	Panic disorder
B	F	25-30	Single	Higher	Internship	PDT (20)	No	GAD
C	F	35-40	Cohabiting	Higher	Employed	PDT (20)	Both	OCD; GAD
D	M	35-40	Single	Secondary	Interrupted	PDT (20)	No	None
E	F	55-60	Married	Secondary	Employed	PDT (20)	No	OCD; Pain D.; ED.
F	M	50-55	Married	Higher	Unemployed	PDT (20)	Therapy	None
G	F	50-55	Divorced	Higher	Housewife	PDT (7)	Meds	Somatization D., BDD
H	F	30-35	Single	Secondary	Unemployed	PDT (20)	Both	Agoraphobia; BDD
I	M	20-25	Single	Secondary	Student	PDT (20)	No	Panic disorder; GAD
J	F	25-30	Single	Secondary	Interrupted	PDT (20)	Both	Social Phobia; GAD; ED
K	M	25-30	Single	Higher	Unemployed	CBT (20)	Therapy	None
L	M	30-35	Single	Secondary	Unemployed	CBT (12)	Both	Specific phobia; OCD
M	F	20-25	Cohabiting	Secondary	Employed	CBT (6)	Both	None
N	F	25-30	Single	Higher	Employed	CBT (20)	Therapy	PTSD
O	F	20-25	Cohabiting	Higher	Student	CBT (17)	Both	GAD
P	F	50-55	Divorced	Secondary	Employed	CBT (20)	Both	None
Q	M	35-40	Cohabiting	Higher	Employed	CBT (8)	No	None
R	M	40-45	Single	Higher	Employed	CBT (20)	Both	Panic Disorder
S	F	25-30	Cohabiting	Higher	Employed	CBT (20)	Therapy	Panic D., Agoraphobia; Social phobia; OCD; GAD; PTSD; Hypochondrias

Note. Information as indicated prior to therapy. To safeguard participants' anonymity, no exact ages are mentioned in the table. M/F: male/female. 'Cohabiting': living together with romantic partner. 'Higher education': college or university degree. 'Interrupted employment' (i.e., temporarily): e.g., due to sick leave. 'Prior care': previous psychotherapy or medication (i.e., antidepressants or other psychopharmaceutic treatment). 'Both': medication and psychotherapy. GAD: generalized anxiety disorder; OCD: obsessive compulsive disorder; ED: eating disorder; BDD: body dysmorphic disorder; PTSD: post-traumatic stress disorder.

Grounded Theory Analysis

Grounded Theory can be described as an explorative and interpretative qualitative research method, aimed at the construction of new theories or rationales grounded in data (in our case patient interviews) (Charmaz, 2014; Fassinger, 2005). Using this method, a tentative conceptual

model of non-improvement that comprises patients' experienced changes and explanatory factors was created Characteristic of grounded theory, several stages of analysis were completed in a cyclic manner before arriving at the final conceptual model (Mortelmans, 2013). This form of inquiry enabled the exploration of the phenomenon of non-improvement in the participants' terminology and to identify themes in the data in a bottom-up manner. As the interviews were conducted in the context of a larger study, the interview questions were not altered throughout the data gathering process as is often the case in grounded theory analysis.

Prior to the actual coding of the interview transcripts, the first author wrote a vignette about every participant that included demographic information, treatment duration, pre-post outcome scores and a summary of the most important themes addressed in the interview. The vignettes were used to get an initial idea of the individual cases in the sample prior to the analysis. During later stages, the first author repeatedly reread the vignettes to validate the constructed model and conclusions with the individual cases. The interview transcripts were subsequently analysed by the first author in dialogue with the third author; the second author functioned as an auditor throughout the process (Hill, Thompson, & Williams, 1997).

Open coding. Open coding is defined as 'the analytic process through which concepts are identified and their properties and dimensions are discovered' (Strauss & Corbin, 1990, p. 101). In this phase, the interviews were first read and reread to identify relevant parts of the interviews relating to the research questions (cf. selecting meaning units; Giorgi & Giorgi, 2003). Labels were attached to certain parts of the text, differentiating experienced changes/remaining issues from therapy factors and other mentioned influences. Non-relevant parts, i.e., not dealing with the topic of well-being, experienced changes or therapy, for instance, were omitted. In order to prevent relevant information from being omitted during the coding process, this work was first conducted on printed versions of the interview transcripts and repeated in the Nvivo software package. This phase resulted in a first list of codes that were

formulated with the intent to remain close to the narrative of patients. A first rough classification was made between the various codes (i.e., experienced changes, remaining issues, therapy effects, social context). They were discussed between the first and third author and altered until consensus was reached.

Axial coding. Axial coding can be summarized as 'the process of relating categories to their subcategories termed "axial" because coding occurs around the axis of a category, linking categories at the level of properties and dimensions.' (Strauss & Corbin, 1990, p. 123). In this phase, the various codes were further divided into subcategories in order to refine the first initial classification of codes. In dialogue between the first and third author, the resulting codes were thematically connected and where needed rephrased. At the end of this phase, the first author looked for visual images and metaphors that could help to grasp the central categories and mechanisms emerging from the narratives of the patients (e.g., "stuck in a maze"; "impasse"). These were further developed and refined in the next phase.

Selective coding. Selective coding comprises 'the process of integration and refining the theory' (Strauss & Corbin, 1990, p. 143). In this phase the theory was cultivated by creating a core category and building other categories around it. In discussion with the third author, this theory was refined. The second author audited the selection of the core and subcategories by asking critical questions regarding the rationale behind the extracted central mechanism. At the end of this phase a set of subcategories was created based on the entire nonimproved sample. Subsequently, we looked into the frequencies of the different categories represented in the CBT and PDT group in order to unravel therapy-related differences. These were included in a detailed table. After finalizing the theory, adequate and valuable phrases were chosen to describe the categories and informative quotes were selected to illustrate the various categories and their interrelations and influences. Patients were given a letter of the alphabet to anonymise text fragments (i.e., from A to S).

Credibility. Credibility checks were held at several stages of the analysis. At the end of every interview, patients were asked whether they wanted to add further information that had not been addressed in the interview. During the analysis, we tried to remain transparent about the entire process (Stiles, 1993) and we acknowledge the influence of the perspective and background of the researchers. The researchers' personal interest in patients' idiosyncratic perspective for instance instructed the focus of the study and analysis. Potential consequences of implicit guiding assumptions were controlled as much as possible by making this idiosyncratic focus central to our study (Creswell & Miller, 2000). We furthermore departed from the assumption that "non-improvement" can also include changes, therefore this was explicitly integrated in our research questions. We worked in a systematic manner to form conclusions and interpretations (Stiles, 1993) and attempted to stay open for any information coming from the narratives throughout the entire process. The analysis aimed at outlining macro-processes in psychotherapy, i.e., examining a wide angle rather than micro-processes (e.g., specific therapeutic effects) and investigated the subjective experience of several different participants (i.e., between-case variation) (Denzin & Lincoln, 2005). In line with our research aim to investigate therapy and outcome in a broader context, the analysis and interpretation of patients' narratives were conducted using a contextual perspective that departs from the assumption that the broader social context influences how patients give meaning to their experiences (Boyatzis, 1998). Triangulation among researchers, several interviews, and quantitative and qualitative indications of outcome were applied to gain different perspectives on the issue. The ultimate themes were formed by asking critical questions regarding codes and categories (Mortelmans, 2011).

Results

In this section, we will present the quantitative pre-post outcome data and qualitative analysis of patients' experiences respectively. Interpretations of the quantitative and qualitative findings

will be described separately; broader integrative conclusions and implications will be presented in the discussion.

Descriptive Pre-Post Outcome Scores on the Beck Depression Inventory

Table 2 summarizes the average score on the BDI-II-NL (Beck et al., 1996; van der Does, 2002) before and after treatment, the standard deviations (SD) and range in scores (i.e., minimum and maximum score) for the entire nonimproved sample, the PDT and CBT group.

Table 2. Pre-Post Outcome scores on the BDI-II-NL

Total score (BDI-II-NL)	**All *n*=19**		**PDT *n*=11**		**CBT *n*=8**	
	M	SD (range)	M	SD (range)	M	SD (range)
Start therapy	30	5.3 (22-42)	31	10.2 (24-36)	29	6.2 (22-42)
End therapy	30	7.8 (18-46)	29	7.4 (20-44)	30	8.3 (18-46)

Note. Meaning of scores on the BDI-II-NL (van der Does, 2002): 0-13: minimal depression, 14-19: mild depression, 20-28: moderate depression, 29-63 severe depression. Cut-off clinical range: 11.3.

Both at the start and end of therapy, the non-improved patient group is characterized by a wide range in scores. At the start of treatment, patients scores varied between moderate depression (n=10) and severe depression (n=9) (cf. van der Does, 2002). At treatment termination, one patient scored mildly depressed, nine scored moderately depressed and nine others scored severely depressed. The average score both before (30) and after (30) therapy indicate severe depression for the total sample, although at the borderline of moderate depression. All patients remained in the clinical range and did not change reliably in scores compared to the start of treatment.

Conceptual Model of Non-improvement from Depressed Patients' Perspective

The grounded theory analysis of nonimproved depressed patients' narratives resulted in the core category Stuck between "knowing versus doing". Around this core category, a model was constructed consisting of ten subcategories that help to explain this core concept. The

subcategories are divided into the changes and remaining issues patients mentioned and the positive and negative influences patients ascribed these changes/remaining issues to. These influences are referred to as "explanatory factors" and specified as "facilitating factors" and "impeding factors". Figure 2 depicts the conceptual model: The left part of the model comprises positive changes and facilitating factors, the right part of the model shows remaining issues and impeding factors. Table 3 summarizes all core and subcategories in more detail for the entire nonimproved sample, the PDT and CBT group. The frequencies of patients contributing to each category were added.

Figure 2. Conceptual model of non-improvement from depressed patients' perspective

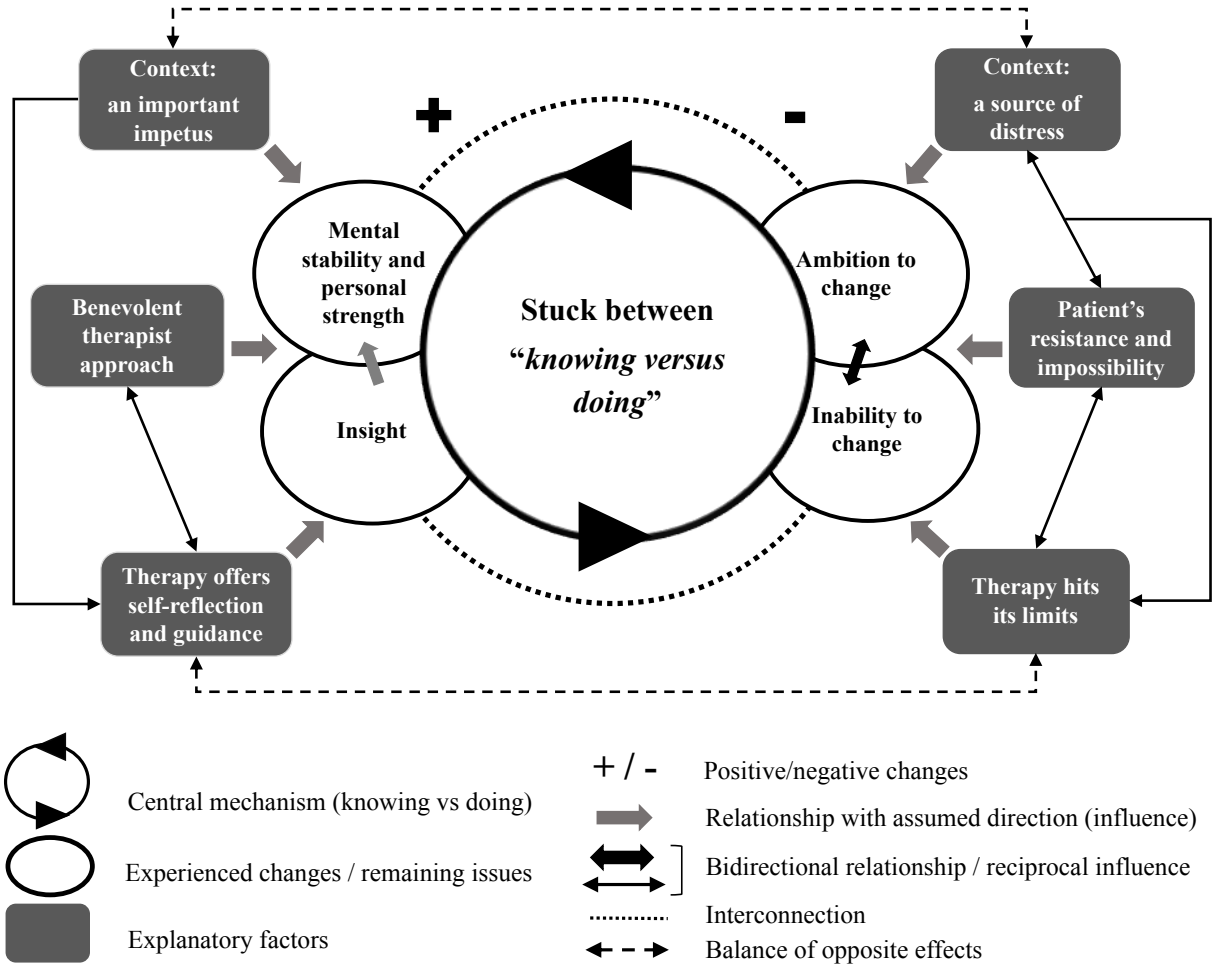


Table 3. Taxonomy of non-improvement based on depressed patients' perspective

	All	PDT	CBT
Core category: *Stuck between "knowing versus doing"*	*n*=19	*n*=11	*n*=8
Positive changes and facilitating factors	***n***	***n***	***n***
Positive changes	16	9	7
Mental stability and personal strength	14	9	5
Insight	15	9	6
Facilitating factors	18	10	8
Therapy offers self-reflection and guidance	14	10	4
Talking and letting it all out	6	5	1
Reflection leading to insight	11	7	4
Provided insight and practical help	12	7	5
Benevolent therapist approach	17	9	8
The right questions	9	8	7
Good relationship	16	9	7
The context as important impetus	11	6	5
Remaining issues and impeding factors	***n***	***n***	***n***
Remaining issues	18	11	7
Ambition to change	13	9	4
Inability to change	13	8	5
Impeding factors	19	11	8
Therapy hits its limits	19	11	8
Something missing	12	8	4
Mismatch and doubt	17	10	7
The patient's:	18	11	7
Resistance	16	9	7
Impossibility	12	9	6
The context as source of distress	16	9	7

Note. All patients mentioned multiple experienced changes, remaining issues and explanatory factors. The indicated frequencies near the categories show this multidimensionality and subcategories thus do not add up. The maximum frequency per category is 19 for the total sample, 11 in PDT and 8 in CBT.

Core category: Stuck between "knowing versus doing". The central mechanism for understanding non-improvement from the patients' perspective is "knowing versus doing": A feeling of having acquired certain changes yet being unable to go a step further, or to know what the problem is but feeling an incapacity to do deal with it. In general, patients wanted to move forward but felt unable to. Some patients stated this literally: 'Rationally I know what my

problems are or what I should do, but there is just nothing changing.' At the same time, the core category captures the effects of therapy that on the one hand facilitated patients' self-understanding and mental stability, but on the other have not been able to overcome certain barriers. A plus (for positive changes and positive influences) and minus (for remaining issues and negative influences) seem to cancel out each other, resulting in a stalemate.

> 'I have learned a lot, I have gained many insights. (…) I am not as despondent anymore, but if I say that "not that much has changed" I mean, I still have difficult periods and a few fundamental problems, which I do understand better now, are not really solved yet, or maybe they are not easy to solve. So, I know much more, I have improved on the level of knowing, but not so much on the more practical level.' (Patient C., CBT)

Positive changes. The overarching experience of being stuck does not imply that patients have not experienced changes at all. Two themes resulted from the analysis showing that throughout the process of therapy, patients have grown mental stability and strength (n=14) and have gained more insight (n=15). Moreover, these changes seem interconnected, as increased understanding was said to have influenced patients' personal strength. In the model in Figure 2, this is indicated by a dotted line (showing interconnection) and an arrow from 'Insight' to 'Mental stability and personal strength'.

> 'In general, I feel much stronger, mentally. I have gained many insights in therapy. (…) It gave me peace of mind and recognition that okay, my thoughts, experience and things I long for are not that strange.' (Patient J., PDT)

Mental stability and personal strength. Increased mental stability (n=11) consisted of two subthemes: A more positive state of mind (n=9) and the ability to accept and let things go (n=6). Patients had learned to deal with certain situations, they felt less emotionally overwhelmed and believed they could handle challenges better. Patients' personal strength (n=11) consisted of an increased self-confidence (n=11) and being more vigorous (n=9). Patients felt more active, dared to socially interact and felt less anxious and insecure.

> Interviewer: 'So you say you are mentally not dispirited anymore. Can you explain the difference with before?'
> Patient: 'I think that before I was really struggling with myself, my self-image and now I feel more balanced. I can take a distance. The fights [with partner] are still intense, but I relate it to the relationship now, I've stopped blaming myself.' (Patient Q., CBT).

Insight. Increased insight was described on three levels. Patients were able to see things from a different perspective (n=9), understood themselves better (n=8) and had gained insight into the reasons why they experienced difficulties (n=8).

> 'Therapy made me reflect upon things and gave me some different ideas about situations that were clear to me but might not have been that clear after all or that needed to be looked at from a different perspective.' (Patient A., PDT).

> 'I have learned more about myself. I already knew I was stubborn, but we talked about it a lot [in therapy], about how I set my standards very high and I don't accept help from anyone and that, by this, I make life very difficult for myself.' (Patient H., PDT).

Facilitating factors. The facilitating factors contributing to these positive changes according to patients include the role of therapy, the therapist and the patients' social and professional context. As indicated in Figure 2 by unidirectional arrows, these features were described as contributing to patients' mental stability and strength as well as insight.

Therapy offers self-reflection and guidance. Patients in CBT and PDT presented slightly different experiences of therapy. These differences seem in line with the specific nature of both types of therapy, given the explorative and expressive style of PDT and the directive approach in CBT. Especially in the PDT group, weekly therapy sessions were considered important for being able to talk, express feelings and thoughts (n=6) and therapy was seen a as a weekly moment of self-reflection (n=11). This seemed to have stimulated both patients' insight (cf. self-reflection) and mental stability and strength (e.g., letting things out; relief).

> 'By saying things out loud in therapy, you start reflecting on them and it becomes a reality that does not only exist in your head. We had one session where the therapist asked some questions about my relationship and because I

> really did not want to answer them, I started wondering, how I really feel in this relationship, I realised that wasn't very good.' (Patient B., PDT).

In the CBT group, similar aspects were mentioned (n=4), but therapy was also valued for actively providing patients with insight and practical help (n=5), making the facilitating role of therapy more guiding than reflective.

> 'She [the therapist] helped me to take certain steps. We tried to come up with means to…, like a priority list for my day, because it is difficult for me to…, although I know I have to do lots of things, I'm not organized.' (Patient R., CBT)

Benevolent therapist approach. Patients in both therapy groups ascribed important changes more to the therapist's way of being than to the therapy form as such. Therefore, we explicitly describe therapy and therapist separately, even though the factors are clearly intertwined. This reciprocity is indicated in the conceptual model (Figure 2) by means of a bidirectional arrow. Firstly, a good therapeutic relationship was described by the majority of patients (n=16) as making it easier for them to talk and open up. More specifically, patients valued that the therapist did not judge but rather encouraged and acknowledged them (n=10). Secondly, the therapist had the skill to ask the right questions (n=15) that stimulated reflection and provided patients with a different perspective. The previously discussed themes 'insight' and 'mental stability' were described as influenced by this approach of the therapist. Although the therapist was mentioned as important in both CBT and PDT, the nature of the experienced role of the therapist's different: In PDT, the therapist was described as stimulating a mental process that was ongoing for the patient, while the CBT therapist was characterized as being an active participant or initiator in the therapy sessions.

> 'His questions, they often appeared so innocent, but when you think about it afterwards you see things from a different perspective. Very subtle because he does not tend to give his own opinion'. (Patient C., PDT).

> 'After a few sessions, my therapist came up with a scheme that summarized my life until now and where my anxiety comes from. I remember I had to cry

> for the first time. I noticed how good it felt to finally be understood. I still believe he can help me in the process of accepting it [certain life events].' (Patient S., CBT).

The context as an important impetus. Besides therapy and the therapist, patients mentioned the influence of their social context as a third facilitating factor. Significant others appeared as an important motivation to do something about problems, to find a job or to keep on going (n=11). For some patients, it had been important to be at home for a while (with sick leave) and to have the time for themselves in their own space (n=10). For others (n=5), work (i.e., the professional context) was an important support mechanism, as it gave them a reason to get up in the morning and structure to their day. The motivating context was therefore considered a facilitating factor for patients' mental stability and personal strength and seemingly a potential stimulus for engaging or continuing treatment.

> 'My son is one of the reasons for starting and continuing the course [i.e., education]. I want to be able to show him something, instead of being an unhappy person. I want to give him something, something positive. That's actually the only valuable thing in my life that's left. I used to be so materialistic, now the only thing that matters is him.' (Patient L., CBT).

Remaining issues. Despite the positive changes in mental stability, personal strength and gained insight, certain issues remained. Feeling stuck was characterized by the wish or ambition to change yet feeling unable to do so. A reciprocal arrow between the categories 'Ambition to change' and 'Inability to change' represents this equilibrium.

Ambition to change. The majority of the patients experienced an ongoing struggle and carried hopes for further change (n=13). These aspects implied things they believed they still lacked or they should work on in the near future, such as tackling self-criticism and self-discipline.

> 'I wish I could have a more positive stance in life, to be able to counter my negative thoughts. I want to have the discipline to get things done, but the hours just slip away, I see the days pass by without getting anything done. That's waisted time to me. I want to get a grip of it.' (Patient I., PDT)

Inability to change. In spite of and seemingly in conflict with the ambition to change is an overall feeling of inability (n=13). Patients felt as if they were running behind on things, they lacked initiative and could not force themselves to move forward.

> 'I'm just not on top of things. I constantly feel rushed, but I lose so much time. Searching for things, not able to finish things because you are too distracted. I just can't seem to get out of it. Things pile up, it seems that for every problem I solve I get two in return.' (Patient P., CBT)

Some patients described an inescapable cycle they seemed to repeat over and over again (n=3). This inability was also reflected in what seemed like an internal conflict (n=7), for instance, being stuck in the struggle between wanting to spend more time with the kids but at the same time aspiring for a professional career. One patient was highly preoccupied with a long-lost dream, which made it impossible to pursue a new goal in life.

Impeding factors. Similar to the factors that help explain positive changes, patients brought up potential reasons for why certain issues remained unchanged. These factors include the limits of therapy, the role of the patient and a negative influence coming from patients' social context.

Therapy hits its limits. Notwithstanding its positive effects, therapy was described as hitting a limit by all patients. More specifically, advise or learned techniques were considered valuable, yet only up to a certain point. Patients stated that they were unable to use the techniques when feeling really bad, or they did not find the time to do so. For others, therapy had progressed too slowly, had not been valuable every session, or it had worked for some aspects but not for others. Therapy hit a limit in two ways: Something was missing in therapy (n=12) and/or the therapy mismatched the patient's needs or expectations at some point (n=17). Noteworthy is the observed difference between the PDT and CBT group. A few patients in PDT were displeased that they had not gained the right tools, were not given any directions (n=4) or stated therapy had a varying impact (sometimes it helped, sometimes it did not) (n=3).

> ‘I expected her [the therapist] to give me good advice on how to deal with my problems. How I can worry less, how I can improve my breathing, just some tips. I must have imagined therapy the wrong way, she did not give me any tips. I’m kind of disappointed. (…) I still don’t understand the purpose of talking about all these things, I often felt worse after the session.’ (Patient G., PDT).

Some patients in the CBT group, on the other hand, stated therapy was too superficial and they needed a more intense form of treatment (n=5).

> ‘The first three to four sessions, you tell your whole life story and all that is said about it is just okay, “you suffer most from the discussions [at home/with your partner] so let’s see how we can handle them”. While I thought okay, I just told you my entire life story, about who I am and how I became who I am, that could have been included in therapy, but I actually felt that it wasn’t at all, we just looked at one segment.’ (Patient Q., CBT).

Moreover, therapy as hitting its limits seems to have contributed to a rather ambivalent attitude towards possible continuation of therapy (n=11). One third of the patients had no further need or motivation to continue psychotherapy (n=8). Some of them believed they had dealt with everything or had gotten everything out of therapy that they could, others had lost hope that therapy could help them or were disappointed about the results. Half of the patients indicated they would continue the same therapy, because they felt committed to the process (n=6) or because they had further specific issues they wished to address (n=4). Others, however, were interested in pursuing a different kind of therapy (n=7).

The patient’s resistance and impossibility. Patients also reflected on their own role and position in therapy. Most patients described feeling a certain resistance towards therapy (n=16). For instance, they did not take therapy very seriously, had difficulties with opening up or were reluctant to do certain exercises. Several patients saw therapy as a task or an investment that asked too much from them (e.g., energy, time, money) (n=6). This rather ambivalent position in therapy was, for instance, described by patient R:

> ‘I was afraid to fail in therapy. (…) I typically start things but can’t manage to continue them. Maybe because (…) when it hasn’t gone well one day, I

> can't let that go. It is all or nothing often, so I was afraid I would not do very well [in therapy] and also, sometimes I put effort into it, but often I was busy doing other things I thought I should be doing, like work.' (Patient R., CBT).

Secondly, many patients were convinced about the fundamental nature of their problems and the impossibility to change (n=15), sometimes referring to their own personality. Moreover, some patients indicated that therapy of twenty sessions is in general too little to solve more fundamental problems. The patient's resistance and idea of impossibility seem to correspond to the perspective on therapy as hitting a limit. In Figure 2 this was indicated by means of a reciprocal arrow, assuming a certain reciprocity between both explanatory factors.

> 'I think I'm quite a different case, I have quite a big tendency towards depression, if I compare myself to other people in my environment who have depression, they get over it after a year, but I think for me this is a bit more difficult, because of my childhood... I have been conditioned to think in a certain way, I think that matters a lot [for the duration and effect of therapy].' (Patient O., CBT).

Despite a mismatch, some patients described they were able to get passed initial resistance and adjusted to therapy, while for others, the limits of therapy, their own resistance or feelings of impossibility lasted throughout the therapy process.

> 'I might have expected therapy to be a bit more practical, but I noticed quite fast that this wasn't the goal of the sessions and I accepted that, I didn't see the talking as a waste of time.' (Patient I., PDT).

> 'So, I think "okay, therapy has ended now and once again I'm nowhere, it did not help, and it only cost me money, a lot of time and energy, and why? For nothing." (…) Of course, I know I did make progress, but I it's hard for me to see it that way. I just think, "it's the umpteenth thing I've tried and what is the use?"' (Patient P., CBT).

The context as a source of distress. Several contextual factors were mentioned as having a negative influence on patients' wellbeing during and after therapy. Firstly, patients' personal context was mentioned as being highly stressful (n=16). Several patients, for instance, encountered a conflict with family members (n=6). These difficulties in relationships were often

considered to facilitate or perpetuate certain problems, and consequently they were believed to have influenced the therapy process and patients' progression. Patient F., for instance, described a critical moment in course of therapy:

> 'There was a crisis [during treatment]. It was when I just started working there [family business], it became too much with how my brother-in-law always got angry with me. I could not handle it. He yelled at me that I was making him bankrupted, I cost a fortune, I don't work well, I'm too slow. At a certain point I switched off my phone and just ran away, I wanted to disappear, commit suicide.' (Patient F., PDT).

Secondly, the professional context appeared as a source of stress or dissatisfaction (n=9). This subtheme contains patients who experienced high amounts of stress due to school-related deadlines, had difficulties adjusting to a corporate culture (e.g., not able to handle the given freedom) or did not experience any fulfilment at work. Finally, many patients mentioned external factors causing distress, like certain events or circumstances (n=12), for instance, dealing with unexplainable physical complaints and an ongoing lawsuit. The reciprocal arrow in the conceptual model between the patient's resistance and impossibility and stressful context indicates that both factors are understood as interacting and influencing the therapy and recovery process.

Discussion

This study investigated the phenomenon of non-improvement in psychotherapy from the perspective of depressed patients in relation to their pre-post outcome scores showing no reliable change. By doing so, we answered the pressing need for further investigation of the phenomenon of negative outcome and the exploration of the relationship between quantitative and qualitative approaches to outcome and treatment evaluation in the field of psychotherapy (McLeod, 2013).

First and foremost, the findings of this study showed that non-improvement as indicated by symptom-based outcome scores did not mean that patients did not experience any changes.

Where a lack of change in outcome scores means a status quo on the level of symptom-severity, the interviews of the patients revealed a more nuanced and complex picture. Central to patients' experience of non-improvement is the mechanism of knowing versus doing. While patients had the ambition to change, they felt unable to overcome certain problems, resulting in a stalemate of knowing what to change but not being able to. Positive changes were offset by substantial remaining issues: Increased mental stability, personal strength and insights were gained, yet these did not result in changes on other levels of patients' lives. From the patients' perspective, "no change" in symptom-based outcome scores seemed to be "not enough" change or a "partial change". The therapy, the therapist, the patient and context facilitated positive changes but at the same time were unable to alter important issues or even impeded patients' progression (resulting in remaining issues). None of these factors can be considered the main or only explanatory reason but must be understood as interacting (cf. Mash & Hunsley, 1993; Werbart et al., 2015). In sum, an equilibrium between a positive and negative pole seems to characterize the depressed patients' experience of non-improvement.

A similar positive-negative balance has been observed by Werbart and colleagues in a study on non-improved patients' experience of psychotherapy (2015). Nonimproved patients perceived their therapy as "spinning one's wheels": Therapy was valued for some aspects but disappointing on others and even though some changes occurred, core difficulties remained. The current study investigated nonimproved patients' experiences of outcome and therapy in a broader context of various potential explanatory factors (i.e., not limited to the effects of therapy). In that sense, the findings of this study and the study of Werbart et al. (2015) can be seen as complementing each other, also because different populations of patients were investigated (i.e., adults and young adults). Notably, the experience of both outcome and therapy are strongly congruent in reflecting a balance between a plus (i.e., positive changes and facilitating factors) and minus (i.e., negative changes/remaining issues and impeding factors).

The positive pole of the resulting conceptual model in this study, including increased mental stability, personal strength and insight, corresponds to findings of other qualitative outcome studies. Mental stability and personal strength relate to what has been described as feelings of empowerment and improved emotional functioning (McElvaney & Timulak, 2013), and more generally, as changes on the level of the self (Timulak & Creaner, 2010). Strikingly and in contrast to findings from studies on patients' experience of positive outcome, our nonimproved sample did not report changes on an interpersonal level (Binder et al., 2009; Nilsson, Svensson, Sandoll, & Clinton, 2007; Timulak & Creaner, 2010); reported positive changes were overall more self-focused. Indeed, we could wonder whether improvement on a symptomatic level enables or coincides with changes on a more interpersonal level. Regarding patients' gained insight, our findings seem partially in contrast to the commonly derived conclusion that insight is an important acquisition for obtaining positive outcome (see the recently published meta-analysis of Jennissen et al., 2018). In this study, increased insight facilitated patients' mental stability and personal strength, but it did not alleviate patients' self-reported symptoms or alter core difficulties. Similarly, Lilliengren and Werbart (2005) found that self-knowledge does not always coincide with changes in underlying problems. Qualitative studies suggest an important role for agency regarding this link between insight and outcome: Rather than gaining insight as such, it is important to gain the capacity to apply or act upon gained insight in daily life (McLeod, 2013)31. Stage models of therapy (see for instance Hill, 2004), state that insight is only valuable to the extent that it leads to action. The absence of this active component could explain the lack of improvement on other levels in our research sample. As already stated by Freud, in order to gain substantial change, a step beyond intellectual insight towards experience might be required (see Bohart, 1993; Castonguay & Hill, 2007). More research on the mechanism of how insight promotes change is, however, warranted.

[31] See McLeod (2013) for a summary of qualitative studies on this topic, Bohart & Wade (2013) on the role of agency, Castonguay & Hill (2007) for an elaboration on the role of insight in psychotherapy.

The helping role of treatment differed depending on the type of therapy. In accordance with the finding of Nilsson and colleagues (2007), patients valued CBT and PDT for different reasons. In our study, therapy provided a moment of self-reflection for patients in the PDT group, while practical help and guidance was valued in the CBT group. Interestingly, while patients in both CBT and PDT mentioned the central role of the therapist, its specific effectuation differed seemingly. In line with the differentiation between the approach of the respective therapies, the PDT therapist was attributed a rather subtle though powerful technique stimulating reflection in patients. The CBT therapist, on the other hand, was considered an active participant in treatment who offered patients insight via tools such as schematic overviews. A good therapeutic relationship was one of the most important common factors in psychotherapy (Lambert & Barley, 2001) mentioned by the majority of the patients in our study, similar to other qualitative findings (Levitt et al., 2016).

Nonetheless, all patients stated therapy hit a certain limit. Again, in line with the observation of Nilsson et al. (2007), both types of therapy were criticized on a different basis: In our sample, dissatisfied patients criticized CBT for being too superficial while PDT was criticized for not offering the right tools or direction. The latter corresponds to findings from the study of Lilliengren and Werbart (2005), in which patients experienced similar disappointments in psychoanalytic therapy (e.g., wanting a more active therapist, guidance, feedback and advice). A possible mismatch between certain patients' needs or expectations and the type of therapy supports the increasing emphasis in research to explore which type of therapy works best for whom and focus on the tailoring of treatment to patients' transdiagnostic characteristics (Norcross & Wampold, 2011).

Beyond therapy hitting limits, patients in this study mentioned explicitly that they themselves encountered a certain resistance or hit their own limits and limitations. The patient's in-therapy behaviour, such as client involvement and motivation, is the single most important

predictor of outcome. Patient motivation has moreover been linked to expectations and hopefulness: Patients who do not believe they can change, and who feel hopeless, may have less motivation to participate in therapy (Bohart & Wade, 2013). Accordingly, the participation in the therapy process seemed rather ambivalent in our sample. Notably, individual differences were observed: Some patients were able to get passed initial doubt about the therapy approach and their own ideas of impossibilities, while others did not. Although not mentioned by patients themselves, it is important to consider that many patients in the sample presented with one or more comorbid disorders, in most cases some kind of anxiety disorder at the start of therapy. Previous research has shown comorbidity in general predicts worse outcome (see Lambert, 2013).

Finally, our findings revealed the therapy process was intertwined with influences from outside the therapy room. Patients' personal context was both considered an important motivation as well as a large source of distress. Again, opposite effects facilitated and impeded changes. It has been outlined that the context plays a central role in sustaining involvement in psychotherapy or undermining this effort (Drisko, 2004; Lambert, 1992). The impact of patients' professional context on their well-being mentioned in this study is in line with robust findings on the impact of job satisfaction on mental health (Faragher, Cass, & Cooper, 2005). Whilst most qualitative studies tend to focus specifically on patients' experience of psychotherapy, our study provides a valuable additional element of contextualization.

The resulting negative pole of the conceptual model of non-improvement, including the ambition yet inability to change, shows resemblance to what is considered a central characteristic of experiencing depression: Running behind on things, lacking initiative and motivation (DSM-5; APA, 2013). Feelings of hopelessness and helplessness were unresolved, in line with the remaining average score of severe depression in the sample. However, the question should be posed whether this feeling remained unaltered or rather emerged throughout

the process of therapy, for instance, as a response to a lack of improvement. A pre-post research design, even when including retrospective inquiry of patients' experiences prior to treatment, falls short in answering this question. Longitudinal research that includes patients' experiences at the start of therapy as well as monitors changes throughout the process of therapy is needed (cf. De Smet & Meganck, in press).

The contextual model of psychotherapy as described by Wampold and Imel (2015), offers a valuable framework for interpreting our research findings. In this model, therapy is perceived as a "socially imbedded healing practice" (p. 258) in which the relationship between the therapist and patient is central. According to this model, three pathways lead to change in patients' wellbeing: The first pathway establishes the personal relationship ("real relationship") between patient and therapist, characterized by genuine interest and empathy, the second pathways creates expectations in patients of being able to overcome their difficulties, and the final pathway includes therapy specific features or tasks. Although all three pathways lead to a certain degree of change, central for the therapy to work is that it can engage patients to follow the treatment rationale and overcome personal beliefs and explanations for distress.

Regarding the first pathway, the contextual model assumes that establishing a real relationship with the therapist leads to general well-being rather than symptom reduction (Wampold & Imel, 2015). Correspondingly, we observed an increased mental stability and personal strength while patients remained unchanged on a symptomatic level. Patients' pessimistic expectations regarding improvement moreover remained unaltered and for many it was difficult to adapt to or engage in specific therapy features; the second and third pathway thus seem not (entirely) fulfilled. The dyadic concept of the patient-therapist working alliance (Bordin, 1979) further demonstrates how "the collaborative purposive work" (Hatcher & Barends, 2006, p. 293) was obstructed by this discordance between patient and therapist or

therapy; both by therapy not being able to meet patients' need, as well as by patients' resistance towards the requirements of therapy. This links up to the differentiation between two types of bonds: The work-supporting bond and personal relationship (cf. real relationship). The latter involves affective attachment, liking, trust and respect. The other type of bond is considered necessary for 'the difficult work' in therapy, for instance dealing with affective or painful material or executing assignments like exposure and homework (Bordin, 1979). This distinction helps to understand how, in our study, patients experienced a good therapeutic (or real) relationship but failed to engage in the work-supporting bond. Being at ease in therapy and feeling accepted and understood by the therapist thus seem important, for instance leading to increased well-being (Wampold & Imel, 2015), mental stability and personal strength, though not enough to facilitate further life-changes. Correspondingly, gaining insight or self-understanding as such might not be enough when "the hard work" of dealing with affective material has not been worked through. While the contextual model ascribes most of the responsibility to the therapist, the current study and findings give more weight to the role of the patient and his personal context (Wampold & Imel, 2015).

These findings yield a number of clinically relevant implications. Patients who find themselves stuck between knowing versus doing, may hit a certain limit due to a mismatch with the therapy offer, experiencing personal resistance or encountering difficulties outside of the therapy room. As these implications may be brought about by (idiosyncratic) underlying reasons, it may be worthwhile to take this as a particular clinical focus. In light of the increasing use of routine outcome measures in clinical care (see Boswell, Kraus, Miller, & Lambert, 2015), a lack of changes in symptom severity could indicate any of these reasons and most likely a combination, yet monitoring instruments clearly require further exploration in dialogue with the patient. However, signs of non-improvement may not always be visible for the therapist. Studies have

shown that therapists tend to underestimate negative outcome, as patients tend to keep dissatisfaction about therapy to themselves – possibly because they do not want to offend the therapist (McLeod, 2013; Werbart et al., 2015). Therefore, it may be implicated to work on meta-communication in therapy, to avoid or restore possible ruptures in the therapeutic work and relationship (von Below & Werbart, 2012). On the other hand, a well-established therapeutic relationship could change dissatisfaction about therapy into a negotiation, that is, an active focus point in therapy (cf. Wampold & Imel, 2015). In some cases, referring patients to a different approach that is more in line with patients' own rationale may be warranted (Wampold, 2007), as what may work for one patient, might not work for the other (Norcross & Wampold, 2011). Also, the optimal duration of treatment may differ among patients. In the current study, the number of sessions was fixed at twenty, which may have been too little to facilitate changes for some patients (e.g., the average 'good enough level' has been estimated at 26 sessions; Barkham et al., 2006). Moreover, patients showing high levels of resistance in therapy may benefit from a less directive approach (see Beutler, Moleiro & Talebi, 2002, for an overview of the literature on resistance) in which therapy is adjusted to patients' own pace. This is supported by authors who warn that uniform time limits for treatment may not adequately serve individual patients' needs (Baldwin et al., 2009).

This study addresses the critical concern about misrepresentation of patients' outcome by means of standard outcome evaluation and statistical classification (Hill et al., 2013; Kazdin, 2008). First of all, no reliable change in outcome scores seemingly masked the significant changes experienced by patients and does not allow to represent the particular balance between remaining issues and positive changes. Furthermore, considering patients to be a uniform group based on a similar pattern of outcome scores might overlook important individual differences. None of the patients in our sample stated they were cured, although they did vary in the extent

to which they experienced improvement and whether they wanted to continue treatment. In our study, the pre-post changes in outcome scores seem to give a rough preliminary indication of patients functioning, while the patients' narratives show non-improvement is more complex and diverse than can be grasped by a lack of symptom reduction (in line with Zimmerman et al., 2006; 2012). This observation is not surprising in light of the complexity and heterogeneity of depression experiences (Ratcliffe, 2014). It is plausible to assume that recovering from depression is at least equally diverse and layered (cf. von Below, Werbart, & Rehnberg, 2010). Consequently, the findings of this study shed light on the previously voiced question of how negative outcome and non-improvement should be conceptualized. In general, similar to previous research findings, patients' treatment satisfaction and negative outcome did not show a one-on-one correspondence (Werbart et al., 2015); while all patients stated therapy hit a certain limit, a minority was also clearly dissatisfied. Mash and Hunsley (1993) have argued that 'without a guiding theoretical framework for considering failing treatments, the assessment task is daunting, because almost any event in therapy might be construed as a possible indication that treatment is currently failing or is about to fail.' (p. 293). This study shows how this endeavour benefits from a mixed-methods research format that integrates a grounded theory approach. In line with the strengths of grounded theory (Charmaz, 2014; Fassinger, 2005; Mortelmans, 2011), further theory-building research can mean an important contribution here (cf. Stiles, 2015).

Strengths, Limitations and Future Directions

The implications of this study address the well-known gap between academic research and clinical practice (Castonguay, Barkham, Lutz & McAleavy, 2013). RCTs as golden standard research format are limited in providing knowledge that can inform clinical practice (Westen, Novotny & Thompson-Brenner, 2004). The value of integrating qualitative research into this type of rigorous research has therefore been emphasized (Midgley, Ansaldo & Target, 2014).

The current study provides an actual example and informs both clinicians and research on the relationship between outcome scores and patients' experiences of non-improvement. It furthermore builds on the literature of helping and hindering therapy features (Paulson et al., 2001; von Below & Werbart, 2012) by placing the experience of therapy in a broader context of potential explanatory factors as mentioned by patients.

The current study is one of few examining the relationship between quantitative and qualitative outcome evaluation of non-improvement (McElvaney & Timulak, 2013; McLeod, 2013). Focussing on this particular subgroup rather than deteriorated or dissatisfied patients allows for the contribution to a lack of specificity in outcome research and the literature on negative outcome (Lambert, 2011). Research suggests non-improvement, deterioration and patient satisfaction do not fully correspond, although they are often used interchangeably (Lampropoulos, 2011). The current study gives an overall conceptual model of non-improvement and potential explanatory factors. Whether this is, however, representative for nonimproved depressed patients cannot be concluded. Further research should focus on investigating differences and similarities between various groups of outcomes (cf. recovery, improvement, no change and deterioration; Jacobson & Truax, 1991) in order to get a better understanding of the clinical meaningfulness of change from the perspective of patients.

This study contributes to the understanding of non-improvement in psychotherapy and the relationship between quantitative and qualitative outcome evaluation. It cannot, however, answer the question whether outcome scores were representative for every individual patient. The focus of the present study was to provide an overall understanding, (i.e., a conceptual model) of non-improvement relying on a larger group of nonimproved patients. More idiosyncratic information still remains unaddressed and case-study research focussing on individual patients' narratives and outcome scores is warranted (Kazdin, 2011). Similarly, the study cannot offer a fine-grained comparison of the specific effects of CBT and PDT, which

could be further addressed in research on specific factors. The mixed methods research format in our study furthermore explicitly favoured the qualitative data over the quantitative outcome classification as focus of investigation, limiting the quantitative strand to a single, although psychometrically sound and often used, outcome measure. Our selection of patients based on self-reported symptoms nevertheless, had a considerable impact on our findings. With use of other means for categorization, the sample likely would have turned out differently (e.g., using a different measure, multiple measures or relying on patients' satisfaction). Yet, as the use of statistical classification of clinically meaningful outcome (cf. Jacobson & Truax, 1991) is increasingly common in RCTs and standard outcome research at large (De Los Reyes, Kundey & Wang, 2011), this study explicitly aimed to relate the exploration of patients' experiences to the much-used classification tool. Therefore, the aim of the current study was not to address the issue of measurement as such, nor the validation of the specific questionnaire that was used, but to deepen the understanding of outcome that is gained by these much used categories. Our conclusions on the relation between quantitative and qualitative appraisals of outcome can however not be generalized to the entire field of quantitative outcome evaluation that undoubtedly has evolved in the past decades, for instance with an increasing focus on person-centred questionnaires (Elliott et al., 2016). For the purpose of our study, an explanatory sequential design was most suited (Hesse-Biber, 2010). Nevertheless, further research aiming at different approaches to mixing methods and including idiosyncratic quantitative outcome evaluation could contribute greatly to our knowledge on outcome and psychotherapy.

Given the controlled context of our study (as data was collected in the context of a broader RCT), it offers a strong level of control for confounds. For instance, the research sample was characterized by a primary disorder of major depression, outcome was systematically evaluated in all patients and treatments were manualized. A potential threat is therefore, however, the external validity of the findings (Westen et al., 2004). Unlike in naturalistic

studies, patients with more complex and acute psychopathologies were excluded. Nonetheless, all patients in our study showed comorbid disorders in line with clinical reality; for instance, the co-occurrence of major depression and anxiety disorders observed in this study is a robust finding throughout patient groups (cf. Hirschfeld, 2001). The participants in this study resembled a homogenous and local (predominately Caucasian, Flemish) group of patients, however. Specific (e.g., cultural, ethnic) or more diverse groups of patients could be the focus of complementing research. The research findings might also be biased by a selection of patients willing to participate in the study. Moreover, it is known that patients do not easily disclose negative experiences with therapy or with their therapist, and although interviews can enhance this openness (McLeod, 2000), in general, socially desirable answers cannot be excluded (Thurin & Thurin, 2007).

The model of nonimproved outcome must be considered tentatively, and we do not wish to make strong causal claims regarding the effectiveness of treatment or the causal influence of the therapist. In agreement with Strupp and Hadley (1977), we emphasize that the patient perspective is only one perspective on outcome (e.g., in addition to therapist or societal perspectives), and therefore highlights certain elements while neglecting others. This limits the findings of this study, as previous research has shown patient, therapist and observers' perspectives on outcome not always converge and all add valuable insights for clinical practice (Altimir et al., 2010). Nevertheless, integrating in-depth inquiry of patients' narratives in the form of mixed methods research is of considerable value to outcome research and the study of non-improvement. In general, we argue that further research investigating the complex phenomena of outcome and therapy effects should aim at an integration of multiple methods as well as perspectives to grasp the wider picture (McLeod, 2011).

Conclusion

Non-improvement in psychotherapy from the perspective of depressed patients can be understood as being stuck between knowing versus doing, resulting in a stalemate. Patients described both positive changes on the level of insight, mental stability and personal strength. The remaining issues were characterized by an ambition to change but feeling an inability to do so. No change in depression symptoms based on standard pre-post outcome evaluation thus becomes a partial change when considering patients' experience and shows a more complex picture in line with the complexity of experiencing depression. Investigating non-improvement by integrating in-depth analyses of patients' narratives in the form of mixed methods research proves to be of considerable value for understanding (negative) outcome and treatment effects more general.

References

Altimir, C., Krause, M., de la Parra, G., Dagnino, P., Tomicic, A., Valdés, N., … Vilches, O. (2010). Clients', therapists', and observers' agreement on the amount, temporal location, and content of psychotherapeutic change and its relation to outcome. *Psychotherapy Research, 20*(4), 472–487. https://doi.org/10.1080/10503301003705871

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (5th ed.).* Arlington, VA: American Psychiatric Association Publishing.

Baldwin, A. S., Berkeljon, A., Atkins, D. C., Joseph, O. A., & Nielsen, L. S. (2009). Rates of change in naturalistic psychotherapy: Contrasting dose-effect and good-enough level models of change. *Journal of Consulting and Clinical Psychology, 77*(2), 203-211. doi:10.1037/a001523

Barkham, M., Connell, J., Stiles, W. B., Miles, J. N. V., Margison, F., Evans, C., & Mellor-Clark, J. (2006). Dose-effect relations and responsive regulation of treatment duration: The good enough level. *Journal of Consulting and Clinical Psychology, 74*(1), 160–167. https://doi.org/10.1037/0022-006X.74.1.160

Barlow, D. H. (2010). Negative effects from psychological treatments: A perspective. *American Psychologist, 65*(1), 13-20. doi:10.1037/a0015643

Barrett, M. S., Chua, W.-J., Crits-Cristoph, P., Gibbons, M. B., Casiano, D., & Thompson, D. (2008). Early withdrawal from mental health treatment: implications for psychotherapy practice. *Psychotherapy (Chicago Ill), 45*(2), 247–267. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2762228/

Beck A.T., Steer R.A., & Brown G.K. (1996*). Manual for the Beck Depression Inventory II.* San Antonio, TX: Psychological Corporation

Beutler, L. E., Moleiro, C., & Talebi, H. (2002). Resistance in psychotherapy: What conclusions are supported by research. *Journal of Clinical Psychology, 58*(2), 207–217. https://doi.org/10.1002/jclp.1144

Binder, P. E., Holgersen, H., & Nielsen, G. H. (2009). What is a "good outcome" in psychotherapy? A qualitative exploration of former patients' point of view. *Psychotherapy Research,* 1-10.

Blanton, H., & Jaccard, J. (2006). Arbitrary metrics in psychology. *The American Psychologist, 61*(1), 27–41. doi: 10.1037/0003-066X.61.1.62

Bockting, C., & Huibers, M. (2011). Protocollaire behandeling van patiënten met een depressieve stoornis [Protocol treatment for patients with depressive disorder]; In G. Keijsers, A. van Minnen, & K. Hoogduin (Eds), *Protocollaire behandelingen voor volwassenen met psychiatrische klachten (Volume 1) [Protocol treatments for adults with psychiatric complaints]*. Amsterdam: Boom.

Bohart, A. (1993). Experiencing: The basis of psychotherapy. *Journal of Psychotherapy Integration, 3*, 51-67.

Bohart, C. A., & Wade, A. M. (2013). The client in psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behaviour change* (6th Ed.) (pp. 219-257). Hoboken, New Jersey: John Wiley & Sons, Inc.

Bordin, E. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research and Practice, 16,* 252–260. doi:10.1037/h0085885

Bowie, C., McLeod, J., McLeod, J. (2016). 'It was almost like the opposite of what I needed': A qualitative exploration of client experiences of unhelpful therapy. Counselling and *Psychotherapy Research, 16*(2), 79-87.

Boyatzis, R. E. (1998). *Transforming qualitative information: thematic analysis and code development.* London, England: Sage

Braakmann, D. (2015). Historical paths in psychotherapy research. In O. C. G. Gelo, A. Pritz & B. Rieken (Eds.), *Psychotherapy research: foundations, process and outcome* (pp. 39-66). Vienna, Austria: Springer.

Boswell, J.F., Kraus, D.R, Miller, S.D., & Lambert, M.J. (2015) Implementing routine outcome monitoring in clinical practice: Benefits, challenges, and solutions. *Psychotherapy Research, 25*(1), 619, DOI: 10.1080/10503307.2013.817696

Castonguay, L. G., & Hill, C. E. (2007). *Insight in psychotherapy.* Washington: American Psychological Association.

Castonguay, L., Barkham, M., Lutz, W., & McAleavey, A. (2013). Practice-oriented research: Approaches and applications. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behaviour change* (6th Ed.) (pp 85-133). Hoboken, New Jersey: John Wiley & Sons, Inc.

Charmaz, K. (2014). *Constructing grounded theory (2nd ed.).* Thousand Oaks, CA: Sage.

Cooper, M. (2008). *Essential research findings in counselling and psychotherapy: The facts are friendly.* Sage.

Creswell, J. W., & Miller, D. L. (2000). Determining validity in qualitative inquiry. *Theory into practice, 39*(3), 124-130. Retrieved from https://people.ucsc.edu

De Los Reyes, A., Kundey, S. M. A., & Wang, M. (2011). The end of the primary outcome measure: A research agenda for constructing its replacement. *Clinical Psychology Review, 31*(5), 829–838. https://doi.org/10.1016/j.cpr.2011.03.011

Denzin, N. K., & Lincoln, Y. S. (2005*). The sage handbook of qualitative research*. Thousand Oaks, CA: Sage

De Smet, M., & Meganck, R. (2018). Understanding long-term outcome from the patients' perspective: a mixed methods naturalistic study on inpatient psychotherapy. *Psychologica Belgica, 58*(1), 1-21, DOI: https://doi.org/10.5334/pb.432

Drisko, J. (2004). Common factors in psychotherapy outcome: Meta-analytic findings and their implications for practice and research. Families in Society. *The Journal of Contemporary Social Services, 85*(1), 81-90. Retrieved from http://journals.sagepub.com

Elliott, R., Slatick, E., & Urman, M. (2001). Qualitative change process research on psychotherapy: Alternative strategies. In J. Frommer & D. L. Rennie (Eds.), *Qualitative psychotherapy research: Methods and methodology* (pp. 69–111). Lengerich, Germany: Pabst Science.

Elliott, R., Wagner, J., Sales, C. M. D., Rodgers, B., Alves, P., & Cafe, M. J. (2016) Psychometrics of the personal questionnaire: a client-generated outcome measure. *Psychological Assessment, 28* (3). pp. 263-278. doi: 10.1037/pas0000174

Faragher, E. B., Cass, M., & Cooper, C. L. (2005). The relationship between job satisfaction and health: a meta-analysis. *Occupational Environmental Medicine, 62,* 105–112. doi: 10.1136/oem.2002.006734

Fassinger, R. E. (2005). Paradigms, praxis, problems, and promise: Grounded theory in counseling psychology research. *Journal of Counseling Psychology, 52*, 156– 166. doi:10.1037/0022-0167.52.2.156

First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B.W. (2002). *Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition. (SCID-I/P).* New York: Biometrics Research, New York State Psychiatric Institute

Giorgi, A., & Giorgi, B. M. (2003). The descriptive phenomenological psychological method. In P. M. Camic, J. E. Rhodes & L. Yardley (Eds.), Qualitative research in psychology: expending perspectives in methods and design (pp. 243-274). Washington, DC: American Psychological Association

Glaser, B. G., & Strauss, A. L. (1987). *The discovery of grounded theory: Strategies for qualitative research.* Hoboken, New Jersey: Transaction publishers

Hamilton, M. (1967). Development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology, 1967*(6), 278–96. https://doi.org/10.1111/j.2044-8260.1967.tb00530.x

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.

Hesse-Biber, S. N. (2010). *Mixed methods research, merging theory with practice*. Guilford press, New York

Hill, C. (2004). *Helping skills: facilitating exploration, insight, and action* (2nd Ed.). Washington, DC: American Psychological Association.

Hill, C. E., Chui, H., & Baumann, E. (2013). Revisiting and reenvisioning the outcome problem in psychotherapy: An argument to include individualized and qualitative measurement. *Psychotherapy, 50*(1), 68–76

Hill, C. E., Thompson, B. J., & Williams, E. N. (1997). A guide to conducting consensual qualitative research. *The Counseling Psychologist, 25*, 517– 572.

Hirschfeld, R. M. A. (2001). The Comorbidity of Major Depression and Anxiety Disorders: Recognition and Management in Primary Care. *Primary Care Companion to The Journal of Clinical Psychiatry, 3*(6), 244–254. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC181193/

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

Jacobson, N. S., Roberts, L. J., Berns, S. B., & McGlinchey, J. B. (1999). Methods for defining and determining the clinical significance of treatment effects: Description, application, and alternatives. *Journal of Consulting and Clinical Psychology, 67*, 300–307. doi:10.1037/0022-006X.67.3.300

Jennissen, S., Huber, S., Ehrenthal, J. C., Schauenburg, H., & Dinger, U. (2018) Association Between Insight and Outcome of Psychotherapy: Systematic Review and Meta-Analysis. *American Journal of Psychiatry*. DOI: 10.1176/appi.ajp.2018.17080847

Kazdin, A. E. (2001). Almost clinically significant (p<.10): Current measures may only approach clinical significance. *Clinical psychology: Science and practice,* 8(4), 455-462. https://onlinelibrary.wiley.com /doi/pdf/10.1093/clipsy.8.4.455

Kazdin, A. E. (2006). Arbitrary metrics: implications for identifying evidence-based treatments. *The American Psychologist, 61(*1), 42–71.

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

Kazdin, A. E. (2011). *Single-case research designs: methods for clinical and applied settings (2nd Ed.)* Oxford, England: Oxford University Press

Klein, M. J., & Elliott, R. (2006). Client accounts of personal change in process experiential psychotherapy: A methodologically pluralistic approach. *Psychotherapy Research, 16,* 91-105. doi: 10.1080/10503300500090993

Lambert, M. (1992). Implications of outcome research for psychotherapy integration. In J. Norcross & J. Goldstein (Eds.), *Handbook of psychotherapy integration* (pp. 94–129). New York: Basic Books.

Lambert, M. J. (2007). Presidential address: what we have learned from a decade of research aimed at improving psychotherapy outcome in routine care. *Psychotherapy Research, 17*(1), 1-14. https://doi.org/10.1080/10503300601032506

Lambert, M.J. (2011). What have we learned about treatment failure in empirically supported treatments? Some suggestions for practice. *Cognitive and Behavioural Practice, 18,* 413–420. https://doi.org/10.1016/j.cbpra.2011.02.002

Lambert, M.J. (2013). The efficacy and effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behaviour change (6th Ed.)* (pp.169-218). Hoboken, New Jersey: John Wiley & Sons, Inc.

Lambert, M. J., & Barley, D. E. (2001). Research summary on the therapeutic relationship and psychotherapy outcome. *Psychotherapy: Theory, research, practice, training, 38*(4), 357-361. Retrieved from https://www.researchgate.net

Lambert, M. J., Hansen, N. B., & Bauer, S. (2008). Assessing the Clinical Significance of Outcome Results. In A. M. Nezu & C. M. Nezu (Eds.). *Evidence-based outcome research. A Practical Guide to Conducting Randomized Controlled Trials for Psychosocial Interventions* (pp. 359-378)

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. DOI:10.1080/10503300902849483

Lampropoulos, G. K. (2011). Failure in psychotherapy: An introduction. Journal of Clinical Psychology: *In Session, 67*(11), 1093–1095. doi:10.1002/jclp.20858

Leichsenring, F., & Schauenburg, H. (2014) Empirically supported methods of short-term psychodynamic therapy in depression – Towards an evidence-based unified protocol. *Journal of Affective Disorders, 169,* 128–143

Levitt, H. M., Pomerville, A., & Surace, F. I. (2016). A qualitative meta-analysis examining clients' experiences of psychotherapy: A new agenda. *Psychological Bulletin.* http://dx.doi.org/10.1037/bul0000057

Lilliengren, P., & Werbart, A. (2005). A model of therapeutic action grounded in the patients' view of curative and hindering factors in psychoanalytic psychotherapy. Psychotherapy: *Theory, Research, Practice, Training, 42*, 324-339. Doi:10.1037/0033-3204.42.3.324

Luborsky, L. (1984). *Principles of psychoanalytic psychotherapy. A manual for supportive expressive treatment. NY*: Basic Books.

Mash, E. J., & Hunsley, J. (1993). Assessment considerations in the identification of failing psychotherapy: Bringing the negatives out of the darkroom. *Psychological Assessment, 5*(3), 292-301. DOI:10.1037/1040-3590.5.3.292

Mays, D. T., & Franks, C. M. (1985). *Negative outcome in psychotherapy and what to do about it.* New York, NY: Springer.

McLeod, J. (2000). The contribution of qualitative research to evidence-based counselling and psychotherapy. In N. Rowland & S. Goss (Eds), *Evidence-based counselling and psychological therapies: Research and applications* (pp. 111-126). London: Routledge.

McLeod, J. (2011). Qualitative research in counselling and psychotherapy. London, England: Sage.

McLeod, J. (2013). Qualitative research. Methods and contributions. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed.) (pp. 49-84). New York, NY: John Wiley & Sons.

McElvaney, J., & Timulak, L. (2013). Clients' experience of therapy and its outcomes in 'good' and 'poor' outcome psychological therapy in a primary care setting: An exploratory study. *Counselling and Psychotherapy Research, 13*(4), 246–253. https://doi.org/10.1080/14733145.2012.761258

Meganck, R., et al. (2017). The Ghent Psychotherapy Study (GPS) on the differential efficacy of supportive-expressive and cognitive behavioral interventions in dependent and self-critical depressive patients: study protocol for a randomized controlled trial. *Trials, 18*(126), 1-11. DOI:10.1186/s13063-017-1867-x

Midgley, N., Ansaldo, F., & Target, M. (2014). The meaningful assessment of therapy outcomes: Incorporating a qualitative study into a randomized controlled trial evaluating the treatment of adolescent depression. *Psychotherapy, 51*(1), 128-137. Doi:10.1037/a0034179

Mohr, D. C., (1995). Negative outcome in psychotherapy: A critical review. *Clinical Psychology: Science and Practice, 2*(1), 1– 27. doi:10.1111/j.1468-2850.1995.tb00022.x

Mortelmans, D. (2011). *Kwalitatieve analyse met Nvivo.[Qualitative analyses with Nvivo].* Leuven, Belgium: Acco

Mortelmans, D. (2013). *Handboek kwalitatieve onderzoeksmethoden (4th ed) [handbook of qualitative research methods].* Leuven: Acco

Nezu, A.M., Ronan, G.F., Meadows, E.A., & McClure, K.S. (2000). *Practioner's guide to empirically based measures of depressions.* Dordrecht: Kluwer Academic Publishers

Nilsson, T., Svensson, M., Sandell, R., & Clinton, D. (2007). Patients' experiences of change in cognitive-behavioural therapy and psychodynamic therapy: a qualitative comparative study. *Psychotherapy Research, 20*, 37-41.

Norcross, J. C., & Wampold, B. E. (2011). What works for whom: Tailoring psychotherapy to the person. *Journal of Clinical Psychology, 67*(2), 127-132.

Ogles, B. M., Lunnen, K. M., & Bonesteel, K. (2001). Clinical significance: History, application, and current practice. *Clinical Psychology Review, 21*(3), 421-446. https://doi.org/10.1016/S0272-7358(99)00058-6

Paulson, B. L., Everall, R. D., & Stuart, J. (2001). Client perceptions of hindering experiences in counselling. *Counselling and Psychotherapy Research, 1*(1), 53– 61. Doi:10.1080/14733140112331385258

Ratcliffe, M. (2014). *Experiences of depression: A study in phenomenology.* OUP Oxford.

Stiles, W. B. (1993). Quality Control in Qualitative Research. *Clinical psychology Review, 13,* 593-618. doi: 10.1016/0272-7358(93)90048-Q

Stiles, W. B. (2015). Theory-building, enriching, and fact-gathering: Alternative purposes of psychotherapy research. In O. Gelo, A. Pritz, & B. Rieken (Eds.), *Psychotherapy research: General issues, process, and outcome* (pp. 159-179). New York: Springer-Verlag.

Strauss, A., & Corbin, J. (1990). *Basics of qualitative research techniques and procedures for developing grounded theory*. London: Sage Publications.

Strupp, H. H., & Hadley, S. W. (1977). A tripartite model of mental health and therapeutic outcome: With specific reference to negative effects in psychotherapy. *American Psychologist, 32,* 187–196. doi:10.1037/0003- 066X.32.3.187

Timulak, L., & Lietaer, G. (2001). Moments of empowerment: A qualitative analysis of positively experienced episodes in brief person-centered counselling. *Counselling and Psychotherapy Research, 1*(1), 62-73. https://doi.org/10.1080/14733140112331385268

van der Does, A.J.W. (2002). BDI-II-NL. *Handleiding. De Nederlandse versie van de Beck Depression Inventory-2nd edition*. Lisse: Harcourt Test Publishers.

von Below, C., & Werbart, A. (2012). Dissatisfied psychotherapy patients: A tentative conceptual model grounded in the participants' view. *Psychoanalytic Psychotherapy, 26*(3), 211-229. Doi: 10.1080/02668734.2012.709536

von Below, C., Werbart, A., & Rehnberg, S. (2010). Experiences of overcoming depression in young adults in psychoanalytic psychotherapy. *European Journal of Psychotherapy and Counselling, 12*(2), 129-147. Doi: 10.1080/13642537.2010.482745

Wampold, B.E. (2007). Psychotherapy: the humanistic (and effective) treatment. *American Psychologist, 62*(8), 855-73. DOI:10.1037/0003-066X.62.8.857

Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work.* New York: Routledge.

Westen, D., Novotny, C. M., & Thompson-Brenner, H. (2004). The Empirical Status of Empirically Supported Psychotherapies: Assumptions, Findings, and Reporting in Controlled Clinical Trials. *Psychological Bulletin, 130*(4), 631–663. doi: 10.1037/0033-2909.130.4.631

Werbart, A., von Below, C., Brun, J., & Gunnarsdottir, H. (2015). “Spinning one’s wheels”: Nonimproved patients view their psychotherapy. *Psychotherapy Research, 25*(5), 546-564.

Wierzbicki, M., & Pekarik, G. (1993). A meta-analysis of psychotherapy dropout. Professional Psychology: *Research and Practice, 24*(2), 190–195. http://dx.doi.org/10.1037/0735-7028.24.2.190

World Health Organization. (2017). *Depression Fact Sheet.* Retrieved from http://www.who.int/mediacentre/factsheets/fs369/en/

Zimmerman, M., McGlinchey, J. B., Posternak, M. A., Friedman, M., Attiullah, N., & Boerescu, D. (2006). How should remission from depression be defined? The depressed patient’s perspective. *American Journal of Psychiatry, 163*(1), 148-150. Retrieved from http://ajp.psychiatryonline.org

Zimmerman, M., Martinez, J. A., Attiullah, N., Friedman, M., Toba, C., Boerescu, D. A., & Rahgeb, M. (2012). Why do some depressed outpatients who are in remission according to the Hamilton Depression Rating Scale not consider themselves to be in remission? *Journal of Clinical Psychiatry, 73*(6), 790-795. https://doi.org/10.1002/da.21987

VI. Making sense of deterioration in psychotherapy: An interpretative phenomenological analysis of patients' perspectives

Abstract

Aim: This study scrutinizes the meaning of deterioration in psychotherapy beyond the widely used statistical definition of reliably symptom increase pre-to-post treatment. Methods: An interpretative phenomenological analysis (IPA) was conducted to thoroughly investigate how deteriorated depressed patients experience their singular outcomes and therapy processes, with specific aim to elucidate both intra- and extra-therapeutic influencing factors. In a Randomized Controlled Study on the treatment of Major Depression, three patients showing reliable increase in symptom severity on the BDI-II pre-to-post therapy were selected for the IPA study: one patient followed CBT, two patients engaged in psychodynamic therapy. Individual interviews were conducted pre-, peri- and post-treatment. Results: Cross-case analysis of patients' post-therapy experiences showed similarities on three domains: (1) enduring complaints, experienced as uncontrollable; (2) non-internalized insights and unanswered questions; and (3) enduring (mild to severe) interpersonal difficulties. Within-case idiosyncratic differences revealed that the statistical classification of "deterioration" not necessarily corresponds to a "deteriorated experience", nor univocally indicates unwanted therapy effects. Our findings point at the influences of the patient's agency in the process, a discrepancy between patients' expectations and the therapy offer, the therapeutic relationship, interpersonal difficulties and contextual influences. Conclusion: The meaning of symptomatic deterioration should be interpreted within a patient's idiosyncratic context. The multi-faceted nature of deterioration requires further research to rely on multiple perspectives, methods and theoretical approaches.[32]

[32] De Smet, M.M., Cornelis, S., Acke, E., Reitske, M., Truijens, F., & Desmet, M. (2019) Making Sense of Deterioration in Psychotherapy: An Interpretative Phenomenological Analysis of Depressed Patients' Perspectives (manuscript forthcoming)

Introduction

Large-scale systematic psychotherapy studies have repeatedly shown that the majority of patients benefits from psychotherapy, yet, at the same time implying that a significant minority does not. In effect, up to 10% of patients are shown to deteriorate in therapy (Lambert, 2011). Despite this substantial percentage, patient deterioration, as well as negative psychotherapy effects have gained little attention in scientific research. Consequently, both elements remain poorly understood phenomena, despite their high clinical relevance (Barlow, 2010). In order to increase the effectiveness of treatment for those patients who are seemingly worse off post-treatment, it is crucial to gain insight into the specific meaning of patient deterioration and the processes underlying this outcome (Mash & Hunsley, 1993). This requires for research to move beyond a nomothetic approach, in which findings are predominantly reported as averaged therapy effects. An in-depth idiosyncratic understanding of individual outcome is warranted, allowing for the identification and understanding of patients that seemingly do not benefit from treatment (Barlow, 2010).

Research on patient deterioration and negative therapy effects is complicated by a lack of clear conceptualization, with various definitions being used in different contexts and often with little consistency (Lambert, 2011; Mash & Hunsley, 1993). Dealing with the lack of clear conceptualization, psychotherapy research relies strongly on the measurement of symptom severity, in which symptom increase is taken as an indication of deterioration (Dimidjian & Hollon, 2010; Ogles, 2013). A consequence of this operationalization is that the question shifts to *how much* symptom increase is clinically meaningful and potentially alarming. This judgement is typically settled by means of statistical definitions of "clinical significance", of which the Jacobson and Truax method (the "JT method", 1991) is the most widely used (Ogles, 2013). The JT method provides an individual patient estimation of the clinical relevance of measured changes and defines deterioration as a reliable increase in symptom severity (Jacobson & Truax, 1991; Wise, 2004).

Although its widespread use and acknowledged contribution to the interpretation of statistical outcome findings, the JT method is criticized for being too conservative, which entails the risk of underestimating patient deterioration (see Wise, 2004). Apart from that, the (symptom-based) outcome questionnaires that are used to measure pre-to-post changes, may not cover what is actually meaningful in patients' daily lives (Kazdin, 2006). Notably, in the case of major depression, research shows that symptom *decrease* does not necessarily cover patients' experiences of improvement (Zimmerman et al., 2012) and that symptom stability does not readily imply an absence of change when patients' perspectives are consulted (De Smet et al., 2019). What 'deterioration' in terms of symptom increase actually means in patients' own experiences and within their broader life context, remains underexplored, as "non-improved" and "deteriorated" patients are often grouped together (e.g., McElvaney & Timulak, 2013; Lazar, 2017). However, "non-response" and "worsening" are believed to represent distinct clinical phenomena that operate via different mechanisms and thus require different clinical needs (Lambert, 2011; Mohr, 1995). The growing implementation of outcome monitoring in routine clinical care – which strongly relies on definitions of clinical significance to signify if patients are "off track" (Lambert, 2015) - makes a thorough understanding of deterioration in light of patients' subjective experiences not only pivotal for research, but also quintessential for clinical practice.

Answering this need, qualitative research on patients' perspectives is increasingly put forward, as it allows investigating psychotherapy processes and outcome in an in-depth and contextualized manner (McLeod, 2013). Several helping and hindering effects of psychotherapy have for instance been extracted from interviews with patients (see Levitt, Pomerville & Surace, 2016), and both patient (e.g., level of disturbance, motivation) and therapist characteristics (e.g., lack of empathy, level of experience) are considered to influence negative treatment outcome (Lambert, 2011; Mohr, 1995). However, the question of how

psychotherapy can "cause" patients to deteriorate remains unclear (Dimidjian & Hollon, 2010). Moreover, research on treatment failure is deemed particularly complicated by the difficulty to disentangle the effects of therapy from external influences (Mash & Hunsley, 1993), which calls for more in-depth research on the underlying mechanisms (Lazar, 2017).

In the present study, we therefore aim to investigate the perspectives of patients who show reliable deterioration in symptoms over the course of psychotherapy using a mixed methods approach. The JT method of categorization is used as a starting point in order to be able to *broaden* this influential framework of outcome evaluation, by scrutinizing it based on patients' narratives (cf. De Smet et al., 2019). Doing so, the present study aims to move beyond the level of description (i.e., reliable increase in symptom scores) and towards a level of in-depth understanding (i.e., patients' subjective experiences). The central aim of the study is to explore how patients who show deterioration at the level of symptom severity describe their outcome and therapy process, by the analysis of patients suffering from major depression. Importantly, we deliberately adopt a broad understanding of "outcome" and "process", considering the possibility of various influences that are typically not limited to the therapeutic process alone (cf. De Smet & Meganck, 2018). As such, specific focus is given to the elucidation of factors of influence, both within and outside therapy (cf. De Smet et al., 2019). In order to make sense of the experiences in this underexplored group of deteriorated patients, an Interpretative Phenomenological Analysis (IPA; Smith, Flower & Larkin, 2009) is executed.

Method

Design

An explanatory sequential mixed-methods study was conducted (Hesse-Biber, 2010). A first phase of the study comprised a quantitative pre-post outcome evaluation of self-reported depression symptoms. In a second phase, patients' interviews were qualitatively analysed.

Systematic comparison and critical integration of the two strands allowed for a better understanding of both the quantitative and qualitative outcome findings.

Setting

This study is based on data from an RCT on the treatment of major depression (for the methodological protocol, see Meganck et al., 2017). Patients in the RCT were recruited via social media and general practitioners in the area of Ghent, Belgium. Patients included in the study qualified for a diagnosis of Major Depressive Disorder and were randomly assigned to time-limited Cognitive Behavioral Therapy (CBT) or Psychodynamic Therapy (PDT). Outcome evolution was measured pre- and post-treatment by means of quantitative self-assessment, and interviews were conducted pre, peri, post treatment to assess experiences of change. The follow-up period of the RCT spans two years and was ongoing at the time of the analyses.

In compliance with APA ethical standards, the study was approved by the Ethical Committee of Ghent University Hospital (Belgium; EC/2015/0085). All participants signed informed consent prior to the start of therapy agreeing to the use of their individual data (e.g., audiotaped interviews, therapy sessions and questionnaires). For the present study, verbatim transcripts of the interviews were anonymized (both patient and therapist details), patients were given a pseudonym and ages and employment are presented descriptively and unspecified. Any information that could lead to the identification of the patient was omitted from the case description.

Participants

All patients in the RCT (see setting) showing reliable deterioration in symptom severity on the BDI-II-NL were included in the present study. Based on the JT method of clinically significant change (1991), reliable deterioration for the BDI-II-NL total score is reached when a person shows a pre-post increase of 9.6 or up (van der Does, 2002). Three cases were retrieved

based on these selection criteria, hereafter referred to as Mark, Ria and Miriam. Albeit the small number of patients that showed deterioration according to this formula, this allowed for an in-depth analysis of rich cases (cf. Patton, 2002) and exploration of the phenomenon of worsening in therapy from patients' subjective experiences, in line with an IPA philosophy (see Smith et al., 2009). In line with the idiosyncratic focus of IPA, we briefly introduce the cases to provide background context (see also Appendix 2).

Case I. Mark entered therapy in his late 50's. He had been feeling depressed for more than ten years. After a period of improvement, he recently got worse again, which he linked to the beginning of a new romantic relationship, i.e., he claimed to suffer from "the burden of expectations" he believed his partner had of him and which he felt unable to satisfy. His main complaints consisted of extreme fatigue and pressure on the head (which he described as "water inside the head, seeing things blurry, almost like a thick mush you have to swim through". He was unemployed at the start of study and diagnosed with Major Depressive Disorder and Pain Disorder (SCID-IV-TR; First, Spitzer, Gibbon & Williams, 2002). In previous treatment he engaged in couple therapy and medication treatment. In the present study, Mark engaged in twenty sessions of PDT and showed a reliable increase of depressive complaints from pre (BDI=37) to post (BDI=48) therapy; both scores indicate a severe level of depression.

Case II. Ria entered therapy in her 50's. She had been suffering from psychological issues ever since adolescence. Tiredness (due to sleep deprivation; sleep apnea) and what she described as "periods of heavy depression" had been persistent throughout her life. She had a history of psychiatric care and hospitalization with an attempt to commit suicide. Her most recent depressive episode (i.e., prior to engaging in the therapy study) occurred after she had been fired from her job. In contrast to the hoped-for free time "to focus on projects of my own", she unexpectedly fell back in desolation. All interviews centered on her professional life and the practical arrangements of "me-time". At the time of the study, Ria was divorced, self-

employed and diagnosed with Major Depression Disorder, Social Phobia and Agoraphobia (SCID-IV-TR; First et al., 2002). Apart from seeing a psychiatrist (for counselling), she engaged in CBT for sleep disorder in the past. In the present study, Ria engaged in twenty sessions of PDT and evolved from moderate (BDI=21) to severe (BDI=40) depression.

Case III. Miriam entered therapy in her late 40s. She had been struggling with finding meaning in life for as long as she could remember. It proved very important for Miriam to be able to participate in the therapy study, as she believed "the university" would guarantee a certain quality for the provided treatment. Her main complaint was the experience of not truly belonging to this world (being "an alien"), and not being able to connect to other people. She explained this by the fact that she was intellectually gifted and overly sensitive. At the start of the study, Miriam is divorced and employed. She was diagnosed with a primary disorder of Major Depression Disorder (SCID-IV-TR; First et al., 2002). She had consulted several psychotherapists before and received antidepressants in the past. In the course of the present study, Miriam ended CBT after the peri interview (i.e., after 8 sessions), showing a reliable increase in scores over that time (BDI=29 to BDI=39); both scores indicate severe depression.

Treatment and Therapists

Mark and Ria engaged in PDT with a 32-year-old female therapist with five years of relevant clinical experience and training in psychoanalytic therapy. Miriam engaged in CBT with a 25-year-old female therapist with three years of relevant clinical experience and clinical training in CBT. Both therapists received a two-day training in the respective (manualized and time-limited) treatment (i.e., Supportive-Expressive Time Limited manual for Major Depressive Disorder by Luborsky, 1984, and Cognitive-Behavioral Protocol for Depression by Bockting and Huibers, 2011). In both therapies, the predefined treatment duration was 16-20 sessions, delivered with the average frequency of one session per week; therapy sessions lasted approximately 45 minutes.

Measures

Beck Depression Inventory. Patients self-reported symptom severity was measured prior to therapy and one week after treatment ended using the Beck Depression Inventory (BDI-II-NL; Beck, Steer, & Brown, 1996; van der Does, 2002). The questionnaire consists of 21 items that are scored on a scale of 0 to 3 and is divided into a cognitive, somatic and affective subscale. A total score between 0-13 indicates minimal depression, 14-19 mild depression, 20-28 moderate depression, 29-63 severe depression. The questionnaire shows good validity and reliability (van der Does, 2002).

Semi-structured interviews. An adjusted version of the *Clinical Diagnostic Interview* (CDI, non-patient version; Westen, 2002) was conducted at intake, which focuses on patients' experiences of complaints, reasons to seek help and previous health care. The CDI furthermore encompasses questions about life history, professional life and interpersonal relations. An adjusted version of the *Client Change Interview* (CCI; Elliott, Slatick & Urman, 2001) was administered both around the eighth session of therapy (peri) and at therapy ending (post). The interview guide was constructed to evoke participants' experiences of therapy, i.e., the changes they believe occurred during therapy, and whether they ascribed these changes to treatment and/or other factors. Patients were further asked about negative changes or further expectations regarding therapy or improvement. All interviews were conducted at the psychology department (Ghent University, Belgium) by independent researchers. Interviews lasted 100 minutes (intake) and 60 minutes (peri and post) on average. Interviews were audiotaped, and verbatim transcripts were analyzed manually (i.e., using paper and pencil).

Interpretative Phenomenological Analysis

IPA is concerned with the lived-experience and meaning making of individuals within their idiosyncratic context and is typically applied to study underexplored phenomena. Following IPA's aim to contextualize participants' narratives from a psychological (theoretical)

perspective (Smith et al., 2009), several specific secondary questions were formulated alongside the central explorative research question of this study (cf. supra). More specifically, we evaluated whether patients' experiences of outcome corresponded to the statistical classification of "deterioration" and aimed to elucidate which helping or hindering factors were described by deteriorated patients (in line with De Smet et al., 2018; 2019).

Following an IPA sequence, four phases were executed per case before moving onto a next case; a final step consisted of a synthesis of all case analyses. We started by analyzing Mark's narratives, investigated Ria's interviews secondly and studied Miriam's case lastly. The changes reported in post interviews were triangulated with peri-interviews; the intake interviews provided extra contextual information. In every initial phase of analysis, the interviews of a case were extensively read while listening to the audio tapes of the interviews. Subsequently, initial notes were written down, containing descriptive, linguistic and conceptual (i.e., more interpretative) features. Emergent themes were developed based on the initial notes and connections were drawn between the different emergent themes, resulting in a tentative conceptual model (which was developed for every case individually, see Appendix 2). In a final step, patterns across cases were identified and discussed, resulting in an integrated conceptual model that offers a visual representation of the findings. The analyses were conducted by the first, second and third author. Each case analysis was initiated by every coder separately and discussed thereafter. The following steps and eventual synthesis were initiated by the first author, presented to the other coders and discussed until agreement was reached through a consensual process (cf. Hill, Thompson & Williams, 1997).

In describing our findings, we adhere to the IPA philosophy of giving voice to the idiosyncratic stories alongside the core themes that emerges from the cross-case comparison (i.e., final step of the analysis). The resulting tentative conceptual model that will be described in the result section (see Figure 1) was thus constructed accordingly: In presenting the findings,

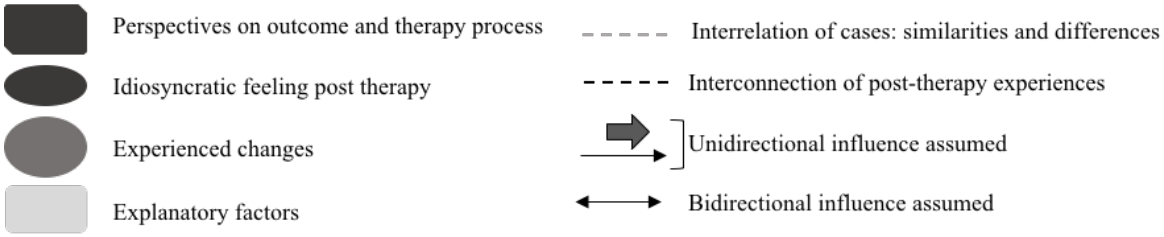
we differentiated *(cross-case) domains of experiences post therapy*, *idiosyncratic (within-case) expressions* and *explanatory factors*.

Quality control. It is important to consider the influence of the researchers' assumptions and choices. For instance, we departed from the possibility that deterioration in symptoms can be accompanied by positive change experiences as well. To tackle personal assumptions as much as possible, we remained transparent about the entire process and attempted to stay open for any information coming from the narratives throughout the entire process. In line with our research aim to investigate therapy and outcome within patients' idiosyncratic context, the analyses and interpretation of patients' narratives were conducted using a contextual perspective. This frame of reference is grounded in the assumption that the broader social context influences how patients give meaning to their experiences (Boyatzis, 1998). We worked in a systematic manner to form conclusions and interpretations (Stiles, 1993). Triangulation among researchers, several interviews per patients, and quantitative and qualitative indications of outcome were applied to gain different perspectives on the issue. If interpretations were made, we provided excerpts to ground our conclusions.

Results

The findings of the IPA are visually represented in the integrative conceptual model depicted in Figure 1. Table 1 summarizes the core themes that will be discussed below. Appendix 2 offers an overview of the central themes and operating mechanisms as found in each separate case.

Figure 1. Integrative Conceptual Model of Deteriorated Patients' Perspectives



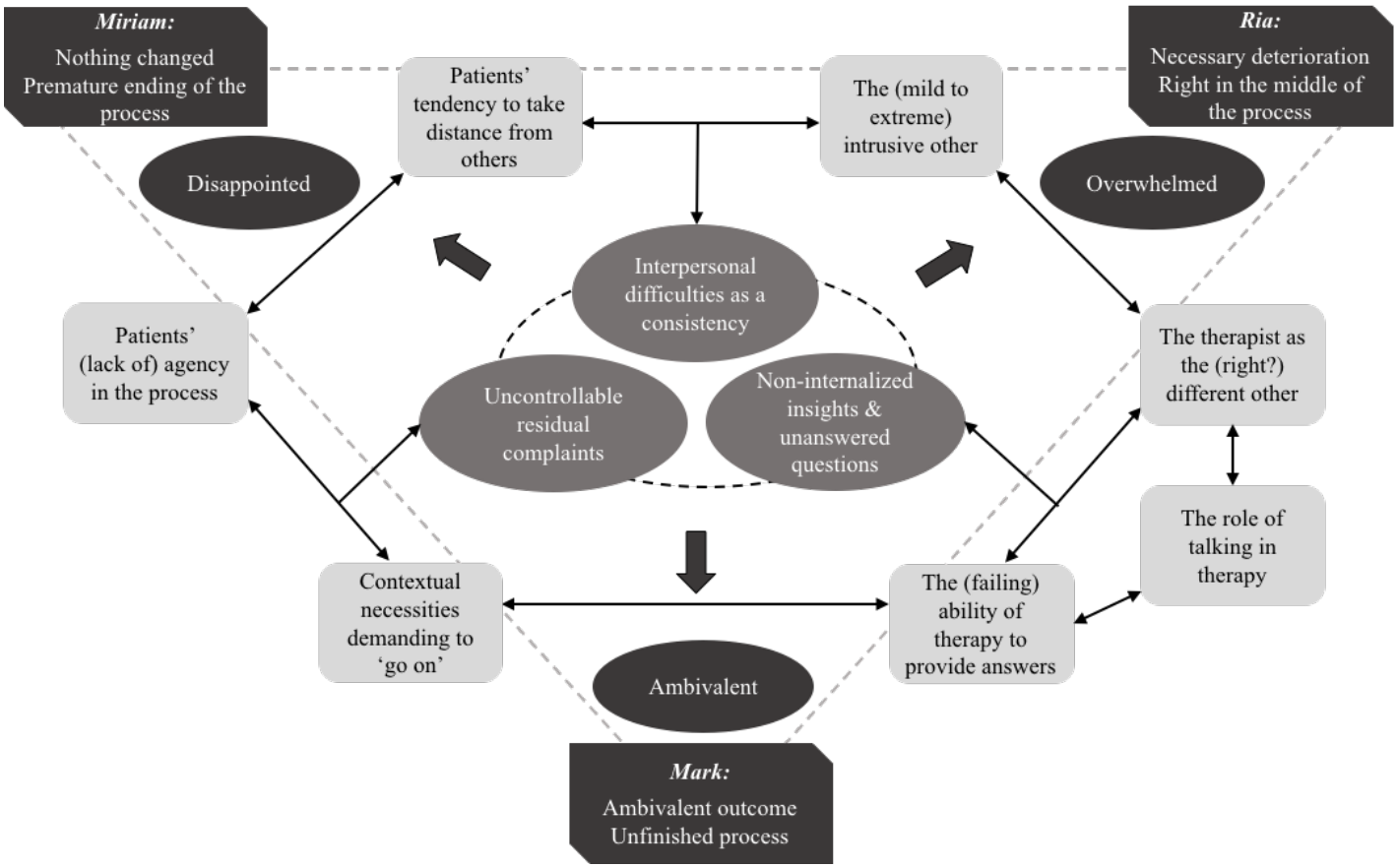


Table 1. *Synthesis of Core Themes*

(Cross-case) Domains of Experiences Post Therapy
Residual complaints experienced as uncontrollable
Non-internalized insights and unanswered questions
Interpersonal difficulties appearing in patients' lives as a persistent consistency
Idiosyncratic (Within-case) Expressions
Mark: Ambivalent – *Is it good, bad, I don't know*
Ria: Overwhelmed – *Things pile up to crisis*
Miriam: Disappointed – *Still can't find meaning to life*
Explanatory Factors
The role of talking: heavy digging or repetitive sharing of information
The (failing) ability of the therapy to provide (workable) answers
The (mild to extremely) intrusive other
Patient's tendency to take a distance from others
Patient's (lack of) agency in the process
Therapist as the (right?) different other
Contextual necessities demanding to "go on" – with or without therapy

(Cross-case) Domains of Experiences Post Therapy

Starting at the heart of the integrative conceptual model (see Figure 1), three central themes capture patients' shared experiences post therapy, which are mutually interconnected (visually illustrated by the dotted line, see Figure 1).

Residual complaints experienced as uncontrollable. All three patients mentioned remaining issues, which they experienced as uncontrollable. For instance, Ria and Mark both expressed recurring debilitating physical complaints; all three cases recounted a sense of tiredness or lack of energy. Mark's physical complaints shifted focus in the course of therapy, however, without a clearly designated meaning: the pressure on his head diminished, while a pressure on his chest – which he related to more panic-like feelings – increased. Importantly, he lacked grip on their occurrence: "it seems to occur periodically, it comes and goes, like a cycle.", he stated. For Ria, the experienced lack of control is illustrated by the fact that circumstances ("even small things") were experienced as *overwhelming.* By the end of the twenty sessions, she felt as if everything "piled up"; she further explained:

> Ria: 'All together my gutter had a leak, there was water in my bedroom, *drip drip drip*, so my sleep, I already have difficulty sleeping. So, on top of that [issues at work]. (…) I tried to arrange the repair of the gutter with the neighbours, but they started arguing. (…) And then later they complained about my dog, because she barks, but of course she's a dog? (…) I mean, I just don't get this "me" mentality?! (…) And all those facts together [becomes too much].'

For Miriam, the uncontrollable nature of her complaints resides in their *incomprehensibility*. She described that she would typically end up in conflict with others, and to feel victimized (*"The intrusive other",* cf. infra), although she had no clue how to avoid these conflicts.

> Miriam: 'The strange thing is that within the same place some people in fact do like me [whilst others detest me]. But what is it then that they find okay? I really don't know. I don't see it, I don't feel it. I actually feel something is not right, but I can't say "oh yes, it's always like that when I act like that". So, I just don't know!'

Non-internalized insights and unanswered questions. Of all three cases, Mark reported the most "insights", yet, described as interpretations by the therapist (e.g., the therapist pointed out that Mark tends to eliminate himself in favour of others). However, his own belief in these insights was expressed with hesitance, for instance: "*apparently* I'm too gullible". Ria on the other hand, especially appreciated the feedback she gained from the therapist.

> Ria: 'I asked her [the therapist] how it is possible that I cannot manage to hold on to a job for two months while the previous person persisted for about six years. And she [the therapist] said, "someone else might take the road of the least resistance and stand all these troubles while you don't take that." I had never looked at it that way.'

Contrastingly, for Miriam, no new insights were gained throughout therapy. She was left with the enigmas of "how to cope with others" and "how to manage social interactions". She stated: "I wished there was someone who could tell me: 'because of *this* specific action, you end up in trouble.'"

Interpersonal difficulties appearing in patients' lives as a persistent consistency. All three cases reported interpersonal difficulties prior to therapy, which, to various extent, had resided or even intensified towards the end of therapy. While Mark indicated that his relational problems were more or less settled, Ria's interpersonal conflicts had only increased (e.g., a dispute at work resulted into her being fired). Miriam's typical pattern of interpersonal conflict remained unaltered. She described how she had been fired consecutively at previous jobs, that she "is used to it by now, it's every year the same story. Everywhere it's the same." She illustrated a conflict as follows:

> Miriam: (…) So at a certain point he just gets angrier and angrier, because I just said something, I replied to his comment. Maybe I'm not supposed to say that, or maybe not in that way. I don't know. But I disobey and then uhm so he got really angry and then at a certain moment I said like "you are not being objective about this" So of course that made his rage even worse.

Idiosyncratic (within-case) expressions

Patients' shared experiences were characterized by an idiosyncratic expression within each case and coincided with a singular perspective on outcome and the therapy process. These case-specific experiences are depicted in the corners of Figure 1.

Mark: Ambivalent outcome – Unfinished process. For Mark, some things had changed during therapy, while others did not. Notably, it was often unclear whether changes were experienced as a good or bad thing. Mark generally expressed a sense of uncertainty, as though he weighed his outcome in a balance.

> Mark: 'Gosh, how am I doing? Good. Depends on how you look at it, of course. I was very down when I started therapy. (…) Well, I mean, the relationship is still going, we are planning to move in together and so on, so that's okay now. But of course, by digging into the past, a lot came to the surface that was very deeply hidden for a long time. Is that a good thing or a bad thing? I don't know.'

Moreover, Mark was convinced he would probably have needed another twenty sessions to "do the thing properly", although he did not initiate to continue treatment (cf. *Contextual necessities demanding to "go on",* infra):

> Mark: 'Gosh, the therapy ending? Well, that's the thing, actually, we both [therapist and patient] sort of felt it was not finished, you know. So, it's actually difficult to really end it right? But yeah, we did follow the [research] procedure, but I mean, it's not easy to end something that's not finished, probably.'

Ria: Necessary deterioration – Right in the middle of the process. Ria, on the other hand, described her situation as a true crisis, a breakdown, as her tolerance to cope with distress decreased ("It feels as if I'm sinking", she mentioned during the peri interview). By the end of therapy, she considered herself in a worsened state. Interestingly, however, she conceived this deterioration as a *necessary phase in order to be able to gain strength again*.

> Interviewer: 'So my first question is, how are you?'
>
> Ria: 'Not good (cries). But I have the feeling that it "has to be this way", in the sense that I felt better before the start of therapy than I do now. (…) However, I believe I really have to experience this now in order to slowly crawl back up

> again. I have to work on what keeps me busy inside, what I struggle with. I feel like I'm at the middle of this process now, which is why I'm a wreck at the moment.'

For Ria, the twenty sessions (as outlined in the context of the study) were only a start; considering herself only halfway through the process, she intended to continue therapy.

> Ria: 'There are a lot of things I need to work on and I'm aware this cannot happen overnight. (…) I think it's normal, so I do see the use of therapy even though I sank deeper than I was before. I feel it was bound to happen someday anyway. (…) And in that case [without therapy] it may have been a lot worse. Now [being in therapy] at least I bear the hope that it will be alright again.'

Miriam: No change – Premature ending of the process. Miriam deliberately ended therapy prior to the intended twenty sessions because she did not experience any change within a reasonable amount of time.

> Miriam: 'Uhm (sighs) nothing has changed. Nothing. On the contrary. (…) I don't expect anything more to come out of therapy. I'm being very honest with you. I think eight times, after eight times you should at least be heading in some sort of direction*.'*

Consequently, Miriam generally expressed disappointment during the interviews; she could still not find a meaning to life, in spite of her efforts to search for help "at the university":

> Miriam: 'I would really like to see it differently, I mean, that's why I participated in this thing [the study]. I thought, maybe at the university, you know, they often have more developed people, maybe they will find the right person [therapist] for me. (…) I feel like, why am I still doing all this effort? I have already put so much energy into it and it just does not work.'

Explanatory Factors

Seven themes (indicated in Figure 1 as *Explanatory factors)* were extracted from the patients' narratives that help to understand the patients' experiences described above (i.e., cross-case experiences and case-specific expressions). The hypothesized operating mechanisms are depicted in Figure 1 and illustrate the mutual interactions between the therapy, therapist, patient, other and broader context.

The role of talking in therapy: Heavy digging or repetitive sharing of information. For all three cases, unanswered questions resided post-therapy (e.g., *"The (failing) ability of therapy to provide answers",* see below). Mark and Ria both described their (psychodynamic) therapy as a process of "bringing (past) things to the surface". For Mark, especially the first therapy sessions were therefore experienced as very heavy and too intense (cf. wishing for a lower frequency).

> Mark: 'It's like a farmer who's digging and spading his field. All sorts of things come to the surface and then some crows and gulls come to pick up the rottenness. (…) Although I did suffer from it, it even pushed me further down at a certain moment, because it's all so long ago. Almost like a previous life, of which you're not consciously aware, or don't want to be aware of. (…) Has everything been said? Probably not. Does everything need to be said? Probably not.'

For Ria, bringing the past to the surface made clear how current difficulties were interwoven with past events (cf. "*Insights"*). Moreover, she stated everything came to the surface, which was clearly audible in the very emotional peri and post interviews (cf. "*Overwhelmed"*). As such, therapy seemed to have taken down Ria's mask, as she described during the intake:

> Ria: 'In the beginning, you can wear a mask and talk and talk and talk, but once you've said everything, and you reach a standstill, you're actually starting off, that's what I felt [during previous therapy], from that point on you start going to your deeper self.'

During the post interview, she described her complaints as "intended symptoms of therapy".

> Ria: You know I have been in therapy before, right? And, the symptoms are just the same, I mean uhm you're in a certain mode, the moment you're busy with daily work and now in therapy it's like everything just comes to the surface.

Contrastingly, for Miriam, talking in therapy (i.e., CBT) was perceived as a means to share information at the beginning of therapy. *Just talking*, however, made therapy move too slow and was at a certain point experienced as mere repetition.

> Miriam: ‘In the beginning [of therapy] you’re just telling your story, and she [the therapist] doesn’t say much, she asks questions, which is logical, but at a certain moment, well I mean, I can always talk, I’m an open book, so the hour always flew by. But it has always been just talking. I just don’t see the use of that, really.’

The (failing) ability of the therapy to provide (workable) answers. Following from *the role of talking*, therapy did not provide concrete or sufficient answers, as opposed to patients’ (initial, but undisclosed) therapy expectations. Both Mark and Ria indicated they missed certain guidance, instructions or feedback. Mark’s hope that the therapist would instruct him to *do this or that, look at this or try that* (as expressed during the peri interview) seemed to have shifted towards the end of therapy, assuming “therapy does not really work like that”.

> Mark: ‘Gosh, yes, I may be too much of a practical person, like, I mean when you go to the doctor’s, he says: “okay take this pill and you can go”. That’s different here of course, so I may have had a different pattern of experience, like okay “this is what you have” I mean, and by now I realize that it doesn’t work like that. You start with the idea “they will tell you what you have et voila (laughs). But that’s not the idea, at least I don’t think it is.’

Ria, on the other hand, anticipated that *a more intense process* was yet to come, and generally believed that the current therapist/therapy was able to help her find the answers she needed. This is in stark contrast to Miriam, who experienced what was offered by the therapist/therapy as insufficient:

> Miriam: ‘The therapist did not tell me anything new and I had hoped to be offered certain things, for instance things I had not thought of myself. I mean, and that sounds very badly, I know, but I can also play my own therapist [when it comes to analyzing myself]. And sometimes she said things and then I was like “no that’s not how it is.’

Therapist as the (right?) different other. The (relation to and person of) the therapist was explicitly mentioned by all three patients and therefore conceptualized as a separate theme. For Mark, the therapist was a *silent listener* who facilitated talking and digging into the past:

> Mark: ‘I always thought like, what do I have to say this time? But then it seems to come spontaneously, she asks a few questions and you pheew, you take off.

> (…) She always managed to find a way to get the conversation going, she could put me in my talking chair.'

Moreover, he described the interaction with the therapist as unique and different from ordinary interpersonal relationships, as she truly listened and never corrected him (in contrast to what people do *in general*). Yet, on the other hand, the therapist's position of silent listener seemingly interacted with Mark's rather uninvolved (or passive) position in therapy (see below "*Patients' (lack of) agency in the process"*).

Ria stated to have found an *ally* in her therapist, whom she experienced as someone who understands her and looks after her needs, in contrast to other people who are "out to get her "and disturb her (cf. "*Intrusive other"*). For Miriam, the therapist personalized *a very kind person* (and in that sense, an exception to Miriam's typical adverse interpersonal relationships), yet *too kind for her*:

> Miriam: So, I think it's a good therapist, but just not for me.' (…) 'Maybe I just haven't found the right therapist yet. Although the therapist, like I said, positive, she is really a nice person, but she cannot help me.

Generally, Miriam indicated to long for a more "unorthodox approach", i.e., being confronted and challenged, which the current therapist had never done.

Patients' tendency to take a distance from others. All three patients reported the explicit wish to oftentimes be left alone (i.e., to be outside the company of others), or to be at ease (i.e., to be away from others' demands and obligations), which especially reflected on Mark's therapeutic process: his wish to be left alone took on a central place in his complaints prior to therapy (cf. relationship demands) and returned as, in order to be able to focus on his future (i.e., "*Contextual necessities demanding to 'go on'",* see below) he needed to leave the past to the past again.

> Mark: 'Do I need extra [therapy]? Gosh, I am actually too busy now with building some kind of a future, uhm moving, living together, uhm ending the company, like finding my way, that I think, everyone, just leave me alone for

> once. I will first do all of this and then, okay we did agree that if I feel like oh what's going on, that I will contact her [the therapist].'

Ria's need to be left alone intensified throughout the course of therapy, as reflected by the increased number of interpersonal conflicts (see above). Initially, Ria felt better, as she – reinforced by the therapist – started doing more things for herself, yet, felt worse again when being *disturbed* or claimed by others (broadly: neighbors, health insurance, obligations).

> Ria: 'Totally unexpected someone called me, you know to sell a product (…). That's really INFUUURIATING! (yells). Do you really have to call me for that?! If I feel like buying the product, I will do so. (…) I would just like to be left alone!'

In Miriam's case, taking distance from others was characterized by a duality. On the one hand, she clearly suffered from feeling *alien* and ending up in conflicts every time, wondering "how I have landed here. (…) I don't have people who think like me. I don't have people I can get along with". On the other, she did not want to be like others and placed herself in an exclusive position. For instance, she did not identify herself with "typical depression" and called herself "an atypical case" requiring "an unorthodox approach*"*.

The (mild to extremely) intrusive other. Patients' need to be left alone seemingly followed from or interacted with experiencing the other as intrusive (e.g., threatening, disturbing, annoying). This often resulted in interpersonal difficulties, as described above; most prominently for Ria and Miriam. As patients explicitly underscored this as an influencing factor, it was captured in a separate theme (i.e., interpersonal difficulties seemed to result from the clash between a need to be alone and experiencing the other as intrusive). Specifically, Ria's discomfort with others amplified towards the end of therapy, by the time she referred to a *maleficent* other who is *out to get her*. It had become increasingly difficult to manage her interpersonal relationships, as distance seemed increasingly necessary, yet was jeopardized.

> Ria: 'Sometimes I think it would be best if I'd just end my life. Like everyone is always shitting on me and like everyone is out to get me. I mean look at my

> job, that certainly hit me hard, they really know what to do to sabotage me, I mean they are really out to get me. What have I done wrong?!'

Similarly, Miriam described a *corrupt* other, who lies, cheats and stabs her in the back. She mentioned several moments in her life when she was bullied and plagued by others.

> Miriam: 'All this backstabbing, it really tears me up inside, how people can lie and be so false. I just don't get it. You see this everywhere, people who stab you in the back… and I am often the victim.'

Patients' (lack of) agency in the process. Patients' interviews also brought up the subject of their own involvement in therapy and the change process. Mark's position appeared to be characterized by a certain *passivity* or *minimal involvement*, strikingly reflected in the fact that he always spoke in general terms rather than applied to himself. Moreover, he often used wordings like *apparently* or *probably* (e.g., "that was difficult, probably"), which we interpreted as expressions of ambivalence or a way of putting his experiences (or himself) at a distance. Conversely, Ria did seem to consider herself as playing an active part in the process, and in that sense disclosed the most involved position in therapy, which may explain why she believed that continuing therapy would help her find what she needs.

> Ria: 'Sometimes I think she [the therapist] could be giving a bit more tips and tricks, so I will discuss this with her on Monday. I really want that because apparently, I really need it.'

In contrast, Miriam stated how she deliberately did not execute the therapist's suggested exercises (e.g., homework), perceiving them as useless. Moreover, she formulated (what was interpreted as) *impossible solutions,* as they almost disabled her to improve, for example, the treatment she needs does not exist:

> Miriam: 'But a therapist will not do that [confront you]. Because that's something you just don't do [as a therapist] that's like too much, but I would really like that. For me that's not an attack, for me that's helping. But yeah, such therapists just do not exist.'

Contextual necessities demanding to 'go on' – with or without therapy. Finally, we observed several concerns in the patients' broader context (i.e., outside of therapy) that showed to influence patients' therapies and change processes (i.e., professional, familial, financial issues). In Mark's case, these contextual urgencies eventually motivated him to focus on building his future (i.e., actively applying for a job, moving in with partner), rather than continuing therapy ("heavy digging in the past"). In contrast, Ria explicitly considered further therapy advisable in order to help her find a job she likes and build her future. Lastly, for Miriam, the intense care for her children meant a motivation to continue her search for therapeutic help elsewhere.

Discussion

This study investigated the phenomenon of deterioration in psychotherapy from the perspective of depressed patients in relation to their pre-post outcome scores showing reliable worsening. By doing so, we answered the need for more idiosyncratic research on the phenomenon of deterioration in psychotherapy and negative therapy effects (Barlow, 2010). An integrative conceptual model is put forward that comprises both (cross-case) shared domains of experiences, as well as (within-case) idiosyncratic perspectives.

Across cases, three domains were identified to characterize patients' post-therapy experiences. Firstly, all three cases described residual complaints that were experienced as uncontrollable (e.g., overwhelming, incomprehensible, periodically recurring). Markedly, patients lacked a feeling of empowerment, i.e., a sense of personal strength or capability to cope with their problems; in great contrast to findings concerning patients who describe positive outcome or helpful therapy effects (Timulak, 2007). Second, whereas Mark and Ria reported to have gained insights throughout therapy, they merely ascribed them to remarks by the therapist, suggesting those interpretations were not internalized (Messer & McWilliams, 2007). Interestingly, a recent meta-analysis by Levitt, et al. (2016) appraised psychotherapy to be

experienced as a “process of change through structuring curiosity and deep engagement in pattern identification and narrative reconstruction” (p.818). In our study, especially Ria seemed to be stimulated to some extent to draw connections, e.g., between past and present events in her life. Contrastingly, and in resemblance with the described cluster “fear of sadness and vulnerability [that] prompts disengagement” (Levitt et al., 2016, p.818), Mark put a stop to an unfinished therapy process, thereby *trusting the painful past to the past again.* A third domain in our model captures the (mild to severe) interpersonal difficulties patients described. Notably, altered or improved interpersonal relationships are oftentimes observed in cases experiencing good outcome (Nilsson, Svensson, Sandell & Clinton, 2007). In our sample, however, only Mark reported improvement (yet exclusively in his romantic relationship), while interpersonal conflicts had even worsened for Ria and remained unchanged in Miriam’s case. As interpersonal difficulties are considered to be at the heart of depression experiences (Joiner, 2000), this observation might be characteristic of this group of patients.

In addition to the cross-case patterns or domains of experience, influential idiosyncratic (i.e., within-case or case-specific) differences were observed. Although all three cases were statistically classified as “reliably deteriorated”, patients appraised their outcome very differently. Only Ria could be considered *deteriorated* when taking her uttered experiences into account, while ambivalence and disappointment about a lack of change was expressed by the others. Moreover, although all three cases indicated the therapy process was not completed at the post interview, they had a different perspective on their therapy process and the point of continuing. Interestingly, Ria perceived deterioration as a necessary phase to gain improvement in the long run. Opposite to the idea of symptom increase as a warning sign for patients being “off track” (Lambert, 2015), Ria felt as being “on track”, given previous successful therapy experiences. Similarly, also Mark described deterioration as part of the therapy process, yet (dissimilarly) this prompted him to disengage from therapy. Miriam ended therapy prematurely,

resembling what has recently been identified as “dissatisfied dropout” (O’keeffe et al., 2019). The cases in this study thus suggest that the statistical category of reliable deterioration in outcome scores, can in fact harbour significantly divergent pictures on the level of patients’ experiences. In other words, considered from a patient perspective, deterioration in outcome scores does not necessarily correspond to a true experience of feeling worsened, nor does it univocally indicate unwanted therapy effects; a finding in line with previous observed discrepancies between patients’ outcome, experiences and therapy satisfaction, cf. De Smet et al., 2019; Werbart et al., 2015).

Making sense of patients’ experiences (i.e., both cross-case experiences and within-case expression), we extracted several mechanisms in mutually interacting factors. First, the role of talking in psychotherapy was experienced as *heavy* due to the fact that *past* events were addressed or *mainly informative* and therefore eventually *needlessly repetitive.* In addition, therapy was experienced as not able to provide concrete answers to lingering questions. Patients seemed to express a mismatch between what they had expected pre-treatment compared to what therapy had brought about. A key feature in patients showing negative outcome (compared to good outcome cases), is often the difficulty of negotiating their therapy (Svanborg et al., 2008), particularly when something seems lacking or when patients are afraid to offend their therapist (Levitt et al., 2016; McLeod, 2013). In the current study, only Ria intended to ask her therapist for more explicit feedback, while Mark and Miriam left this responsibility to the therapist. Thus, in order to tackle potential latent frictions during therapy, it seems advisable for therapists to openly ask for patients’ expectations and discuss the therapeutic framework (e.g., aim of the therapy, manner of working, what is expected of patients’ participation) before the start of therapy. Crawford et al. (2016), for instance, found that patients were less likely to report lasting negative effects if they felt well-informed prior to the start of treatment.

In this regard, patient's agency in the process appeared of pivotal importance, in line with the robust link found between patient in-therapy behaviour and outcome (cf. Bohart & Wade, 2013). Nilsson et al. (2007) observed that among satisfied patients (both in CBT and PDT) the awareness of one's own agency and responsibility in working towards and acquiring change is an important helping factor. For the patients in this study, a lack of agency may partially explain the experienced uncontrollable residual complaints. Remarkably, however, while Ria did not experience any control over external circumstances, she did express an active involvement (and thus sense of control) in the therapy process. In this regard, differences across the patients could be related to the sense of a therapeutic bond, which is considered to be one of the most important common factors in psychotherapy. Ria saw the therapist as an *ally*, together with whom she could overcome the difficult period was in. For Mark, the unique relationship with the therapist deemed especially valuable, although he seemed less inclined to continue the "heavy work" (cf. Gelso, Kivlighan & Martin, 2018, on the role of the "real relationship" in psychotherapy outcome); for Miriam the therapeutic relationship seemed least established. Notably, Mohr (1995) found that particularly patients who exhibit high levels of interpersonal disturbance are at risk for deterioration, as interpersonal patterns might be repeated within the therapeutic relationship. In our study, this was mainly observed in Mark's wish to be "left alone" again.

Essential to the interpretation of our findings, all three cases described to have experienced mental difficulties for a long time, addressing the previous observation that both the degree and chronicity of patients' disturbance can complicate therapy and improvement (Mohr, 1995). As noted by Dimidjian and Hollon (2010), patient deterioration may be a consequence of the natural course of a disorder (which the therapy was unable to alter). In this study, the number of sessions was fixed at twenty, which was experienced as too little to facilitate substantial changes by Mark and Ria. Several authors are therefore critical towards

the implementation of uniform time limits for treatment (see De Geest & Meganck on the influence of time limits in psychotherapy, in press). In line with previous findings, our study also highlights influences outside of the therapy context, which for the cases is this study, mainly involved professional and familial issues (see De Smet et al., 2018; 2019).

Our findings have several implications for research and practice. First, it is generally agreed that deterioration is a complex and multi-faceted phenomenon (Lazar, 2017). Nonetheless, psychotherapy research still strongly relies on symptom evaluation to define the clinical meaningfulness of change (Ogles, 2013). Our findings illustrate how this narrow focus omits important domains in patients' experiences and does not allow grasping the individual or idiosyncratic meaningfulness. Moreover, recent research suggests that similarities between the JT categories might outweigh the differences (cf. Hoffmann et al., 2009; McElvaney & Timulak, 2013), in support of the existing criticism on the dichotomous measurement and understanding of psychological (dys)functioning (e.g., Tingey et al., 1996). Comparing the experiences of the deteriorated patients in this study to those of a sample of non-improved patients (see De Smet et al., 2019), differences seem mainly located at the level of positive changes (e.g., mental stability, self-understanding) and facilitating therapy features (e.g., offering self-reflection and guidance), as they were not explicitely uttered by the deteriorated cases. This suggests that the patients in the present study may indeed represent more severely disturbed cases (e.g., chronical depression) showing worse outcome and negative therapy experiences. It should be noted, however, that the heterogeneity in experiences within both samples, does not allow drawing definite conclusions as to whether it makes sense to cluster patient into separate groups based on quantitative outcome evaluation when considering their own uttered experiences. In other words, we cannot be sure that similar clusters or groups would result from relying on patients' experiences rather than outcome scores when constructing the samples. Therefore, more fine-grained investigation of the relationship

between measured outcome and patients' experiences is warranted and could benefit greatly from case comparison studies (Dattilio, Edwards & Fishman, 2010).

In addition, therapists are generally found to be defective in recognizing patients' (signs of) deterioration, which is why routine outcome monitoring is increasingly advised (Lambert, 2015). However, the so-called ability to recognize deterioration implicitly assumes a clear, common understanding of what deterioration is and how it can be identified. Yet, the field agrees that patient deterioration is complex and hard to grasp, let alone recognized in the course of therapy (Mash & Hunsley, 1993). Moreover, the evaluation of target symptoms (i.e., in our case depressive symptoms) risks to omit intended effects of psychotherapy, as different therapy approaches hold different aims or goals, apart from (or beyond) symptom remission (cf. Lazar, 2017). The use of routine outcome monitoring to signify patients as being "off track" is strongly rooted in the idea that an increase in symptom scores is a warning sign for patient deterioration. However, *what* it specifically alarms for might be less clear, nor univocally provide clear indications as to what therapeutic actions are required (potentially differing from patient to patient). As it remains unclear in what sense a statistical increase in symptom scores reflects negative therapy effects, this should thus be interpreted warily. As such, our findings broaden the interpretation of these signals (i.e., beyond "deterioration" and/or "treatment failure") and emphasize the need to comprehend symptom increase in light of patients' idiosyncratic experiences.

Strengths, Limitations and Future Directions

The current study is one of few examining the phenomenon of deterioration from patients' perspectives in relationship to the often-used statistical definition of reliable increase in symptom scores (cf. McElvaney & Timulak, 2013). Focusing on this particular subgroup rather than non-improved or dissatisfied patients moreover allows for the contribution to a lack of specificity in outcome research and the literature on negative outcome (Lambert, 2011). The

current study gives an integrative conceptual model of deterioration and potential explanatory factors relying on three rich cases. The strength of IPA is the in-depth, idiosyncratic and contextualized study of an underexplored phenomenon. It does not aim nor allow, however, to draw general causal claims, and as such, the presented integrative model must be considered tentatively. Nonetheless, the study sheds light on potential mechanisms that can inform clinical practice, further theory and hypothesis testing. Importantly, the investigated patient perspective resembles only one point of view and not necessarily corresponds to other significant perspectives in the judgement of outcome (e.g., therapist or societal perspectives; Strupp & Hadley, 1977). Further research should aim at investigating deterioration using multiple perspectives, broader: methodologies and theories, to get a better understanding of the multi-faceted nature of the deterioration phenomenon (Lazar, 2017).

Conclusion

Deterioration in psychotherapy from the perspective of depressed patients can be characterized by experiences of uncontrollable residual complaints, unanswered questions and interpersonal difficulties. However, important idiosyncratic differences reveal how the statistical classification of "deterioration" does not necessarily correspond with a *deteriorated experience*, nor does it univocally indicate unwanted therapy effects. As such, the meaning of symptom increase throughout the course of therapy should be interpreted within the framework of each patient's idiosyncratic context. In line with the multi-faceted nature of deterioration, further research should integrate multiple perspectives, rely on multiple methods and consider different theoretical approaches.

References

Barlow, D. H. (2010). Negative effects from psychological treatments: A perspective. *American Psychologist, 65*(1), 13-20. doi:10.1037/a0015643

Beck A.T., Steer R.A., & Brown G.K. (1996). *Manual for the Beck Depression Inventory II.* San Antonio, TX: Psychological Corporation

Bockting, C., & Huibers, M. (2011). Protocollaire behandeling van patiënten met een depressieve stoornis. In G. Keijsers, A. van Minnen, & K. Hoogduin (Eds), *Protocollaire behandelingen voor volwassenen met psychiatrische klachten (Volume 1).* Amsterdam: Boom.

Bohart, C. A., & Wade, A. M. (2013). The client in psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behaviour change* (6th Ed.) (pp. 219-257). Hoboken, New Jersey: John Wiley & Sons, Inc.

Boyatzis, R. E. (1998). *Transforming qualitative information: thematic analysis and code development.* London, England: Sage

De Geest, R. M., & Meganck, R. (in press). How do time limits affect our psychotherapies? A literature review. *Psychologica Belgica.*

De Smet, M. M., & Meganck, R. (2018). Understanding Long-term Outcome from the Patients' Perspective: A Mixed Methods Naturalistic Study on Inpatient Psychotherapy. *Psychologica Belgica, 58*(1), 276. doi: 10.5334/pb.432

De Smet M. M., Meganck, R., Van Nieuwenhove, K., Truijens, F. L., & Desmet, M. (2019) No Change? A Grounded Theory Analysis of Depressed Patients' Perspectives on Non-improvement in Psychotherapy. *Frontiers in Psychology, 10*(588). doi: 10.3389/fpsyg.2019.00588

Crawford, M. J., Thana, L., Farquharson, L., Palmer, L., Hancock, E., Bassett, P., … Parry, G. D. (2016). Patient experience of negative effects of psychological treatment: results of a national survey. *British Journal of Psychiatry*, *208*(03), 260–265. https://doi.org/10.1192/bjp.bp.114.162628

Dattilio, F. M., Edwards, D. J. A., & Fishman, D. B. (2010). Case studies within a mixed methods paradigm: Toward a resolution of the alienation between researcher and practitioner in psychotherapy research. *Psychotherapy: Theory, Research, Practice, Training*, *47*(4), 427–441. https://doi.org/10.1037/a0021181

Dimidjian, S., & Hollon, S. D. (2010). How would we know if psychotherapy were harmful? *American Psychologist*, *65*(1), 21–33. http://dx.doi.org/10.1037/a0017299

Elliott, R., Slatick, E., & Urman, M. (2001). Qualitative change process research on psychotherapy: Alternative strategies. In J. Frommer & D. L. Rennie (Eds.), *Qualitative psychotherapy research: Methods and methodology* (pp. 69–111). Lengerich, Germany: Pabst Science.

First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B.W. (2002). *Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition. (SCID-I/P).* New York: Biometrics Research, New York State Psychiatric Institute

Hill, C. E., Thompson, B. J., & Williams, E. N. (1997). A guide to conducting consensual qualitative research. *The counselling psychologist, 25*(4), 517-572. DOI: 10.1177/0011000097254001

Hoffmann, L. L., Gleave, R. L., Burlingame, G. M., & Jackson, A. P. (2009). Exploring Interactions of Improvers and Deteriorators in the Group Therapy Process: A Qualitative Analysis. *International Journal of Group Psychotherapy; New York*, *59*(2), 179–197. http://dx.doi.org/10.1521/ijgp.2009.59.2.179

Gelso, C. J., Kivlighan, D. M., & Markin, R. D. (2018). The real relationship and its role in psychotherapy outcome: A meta-analysis. *Psychotherapy*, *55*(4), 434–444. http://dx.doi.org/10.1037/pst0000183

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

Joiner, T. E., Jr. (2000). Depression's vicious scree: Self-propagating and erosive processes in depression chronicity. *Clinical Psychology: Science and Practice*, *7*(2), 203–218. doi:10.1093/clipsy/7.2.203

Kazdin, A. E. (2006). Arbitrary metrics: implications for identifying evidence-based treatments. *The American Psychologist*, *61*(1), 42–71.

Lambert, M.J. (2011). What have we learned about treatment failure in empirically supported treatments? Some suggestions for practice. Cognitive and Behavioural Practice, 18, 413–420. https://doi.org/10.1016/j.cbpra.2011.02.002

Lambert, M. J. (2015). Progress feedback and the OQ-system: The past and the future. *Psychotherapy*, *52*(4), 381–390. https://doi.org/10.1037/pst0000027

Lazar, E. (2017). Client Deterioration in Individual Psychotherapy: A Systematic Review. *Theses and Dissertations*. Retrieved from https://knowledge.library.iup.edu/etd/1523

Levitt, H. M., Pomerville, A., & Surace, F. I. (2016). A qualitative meta-analysis examining clients' experiences of psychotherapy: A new agenda. *Psychological Bulletin*. http://dx.doi.org/10.1037/bul0000057

Luborsky, L. (1984). *Principles of psychoanalytic psychotherapy. A manual for supportive expressive treatment*. NY: Basic Books.

Mash, E. J., & Hunsley, J. (1993). Assessment considerations in the identification of failing

psychotherapy: Bringing the negatives out of the darkroom. *Psychological Assessment, 5*(3), 292-301. DOI:10.1037/1040-3590.5.3.292

McLeod, J. (2013). Qualitative research. Methods and contributions. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed.) (pp. 49-84). New York, NY: John Wiley & Sons.

McElvaney, J., & Timulak, L. (2013). Clients' experience of therapy and its outcomes in 'good' and 'poor' outcome psychological therapy in a primary care setting: An exploratory study. *Counselling and Psychotherapy Research*, *13*(4), 246–253. https://doi.org/10.1080/14733145.2012.761258

Meganck, R., et al. (2017). The Ghent Psychotherapy Study (GPS) on the differential efficacy of supportive-expressive and cognitive behavioral interventions in dependent and self-critical depressive patients: study protocol for a randomized controlled trial. *Trials*, *18*(126), 1-11. DOI:10.1186/s13063-017-1867-x

Messer, S.B., & McWilliams, N. (2007). Insight in psychodynamic therapy: Theory and assessment, in L. G. Castonguay & C. E. Hill (Eds) *Insight in psychotherapy*. (pp. 9-30). Washington: American Psychological Association.

Mohr, D. C., (1995). Negative outcome in psychotherapy: A critical review. *Clinical Psychology: Science and Practice, 2*(1), 1– 27. doi:10.1111/j.1468-2850.1995.tb00022.x

Nilsson, T., Svensson, M., Sandell, R., & Clinton, D. (2007). Patients' experiences of change in cognitive-behavioural therapy and psychodynamic therapy: a qualitative comparative study. *Psychotherapy Research*, *20*, 37-41.

O'Keeffe, S., Martin, P., Target, M., & Midgley, N. (2019) 'I Just Stopped Going': A Mixed Methods Investigation Into Types of Therapy Dropout in Adolescents With Depression. *Frontiers in Psychology. 10*(75). doi: 10.3389/fpsyg.2019.00075

Patton, M. (2002) *Qualitative Research and Evaluation Methods* (3rd Ed). Thousand Oaks, CA: Sage.

Smith, J.A., Flower, P., & Larkin, M. (2009). *Interpretative Phenomenological Analysis: Theory, Method and Research*. London: Sage.

Stiles, W. B. (1993). Quality Control in Qualitative Research. *Clinical psychology Review, 13,* 593-618. doi: 10.1016/0272-7358(93)90048-Q

Strupp, H. H., & Hadley, S. W. (1977). A tripartite model of mental health and therapeutic outcome: With specific reference to negative effects in psychotherapy. *American Psychologist, 32,* 187–196. doi:10.1037/0003- 066X.32.3.187

Svanborg, C., Bäärnhielm, S., Åberg Wistedt, A., & Lützen, K. (2008). Helpful and hindering factors for remission in dysthymia and panic disorder at 9-year follow-up: A mixed methods study. *BMC Psychiatry*, *8*(1). https://doi.org/10.1186/1471-244X-8-52

van der Does, A.J.W. (2002). *BDI-II-NL. Handleiding. De Nederlandse versie van de Beck Depression Inventory-2nd edition.* Lisse: Harcourt Test Publishers.

Timulak, L. (2007). Identifying core categories of client-identified impact of helpful events in psychotherapy: A qualitative meta-analysis. *Psychotherapy Research, 17*(3), 310 – 320. DOI: 10.1080/10503300600608116

Tingey, R. C., Lambert, M. J., Burlingame, G. M., & Hansen, N. B. (1996). Assessing clinical significance: Proposed extensions to method. *Psychotherapy Research, 6*, 109–123.

Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work*. New York: Routledge.

Werbart, A., von Below, C., Brun, J., & Gunnarsdottir, H. (2015). "Spinning one's wheels": Nonimproved patients view their psychotherapy. *Psychotherapy Research*, *25*(5), 546-564. Doi: 10.1080/10503307.2014.989291

Westen, D. (2002). *Clinical diagnostic interview manual.* Retrieved from http://www.psychsystems.net/lab.

Wise, E. A. (2004). Methods for Analyzing Psychotherapy Outcomes: A Review of Clinical Significance, Reliable Change, and Recommendations for Future Directions. *Journal of Personality Assessment*, *82*(1), 50–59. https://doi.org/10.1207/s15327752jpa8201_10

Zimmerman, M., Martinez, J. A., Attiullah, N., Friedman, M., Toba, C., Boerescu, D. A., & Rahgeb, M. (2012). Why do some depressed outpatients who are in remission according to the Hamilton Depression Rating Scale not consider themselves to be in remission? *Journal of Clinical Psychiatry*, *73*(6), 790-795. https://doi.org/10.1002/da.21987

VII. General discussion

The present dissertation aimed to open the "black box" of psychotherapy outcome by extending standard statistical outcome evaluation with in-depth analysis of patients' subjective experiences of their therapy outcome and process leading up to this. A multi-layered, mixed methods analysis of psychotherapy outcome was conducted *within* an RCT and comprised three sequential layers of analysis: (1) pre-to-post outcome evaluation interpreted in terms of clinically significant change (Jacobson & Truax, 1991), (2) in-depth analysis of patients' experienced changes, and (3) inquiry of the change processes that have influenced these changes according to patients' perspectives (Carey et al., 2006; Elliott, 2010; Kiesler, 1966). In doing so, we opened up general statistical categories that are dominant in outcome research and looked into the perspectives of the patients on which these categories are build up. Ultimately, this dissertation examined the usefulness of outcome findings as input for policy decisions and organization of clinical care (Kazdin, 2008).

In this final chapter, we juxtapose and integrate the studied patients' reflections on experienced change and change processes across all samples, which allows us to synthesize the understanding of outcome and change processes from a first-person perspective. In a second step, we reflect on the clinical meaningfulness of the influential statistical framework of clinical significance by Jacobson and Truax (1991), in light of patients' own experiences (Atkins et al., 2005). First, the main findings of the separately conducted studies are summarized; secondly, an evaluation of the strengths and limitations of the present research is presented. Implications for research, clinical practice and policy are discussed and recommendations for future avenues are made.

Main study findings

This dissertation aimed to elucidate patients' experienced changes and the change processes underlying or leading up to psychotherapy outcome. In five chapters, we scrutinized the understanding of outcome by examining patients' narratives in the case of long-term, "good" and "poor" outcome. In all the studies presented we reflected on the relationship between standard statistical (pre-to-post) outcome evaluation and patients' experiences and discussed possible implications for the study and conceptualization of outcome in psychotherapy research and daily practice.

Chapter II (Study 1) presented a naturalistic study that originated from a common clinical reality: treatment institutions must prove their effectiveness in order to be funded by government agencies or health care policies. A long-term outcome study was conducted in an inpatient setting; 22 former patients were interviewed five to six years after treatment termination. A data-driven thematic analysis was conducted and resulted in a conceptual model of long-term outcome grounded in patients' narratives. The statistical outcome evaluation on the OQ-45 (Lambert et al., 1996) indicated that the majority of patients had reliably improved during treatment and this until the second follow-up measurement, yet only a few patients reached and remained levels of clinically significant change (i.e., recovery; Jacobson & Truax, 1991). From patients' perspectives, long-term outcome could be understood as an ongoing and fluctuating process and experienced changes were situated on a time line. Concretely, when entering inpatient care, patients experienced reconnection to life and others, something they had lost in what was described "the cold and harsh society". Secondly, patients gained insights and obtained new ways to handle their difficulties (e.g., express emotions), which was experienced as a revelation. Subsequently, changes in the way patients organized their lives took place (during and after the inpatient treatment), which were experienced as life-altering for some. Finally, patients' ideas and expectations regarding recovery and the role of treatment altered (e.g., became more realistic). In terms of helping and hindering factors in and outside

the therapy centre, patients highlighted the interaction between their own position in therapy, the therapy offer and the influence of the external context, which were described as evolving over time.

Importantly, this study points to the *dynamic* nature of influencing factors or change processes. The patients in this study emphasized the importance of being able to *change* during and after the treatment process, which is in contrast to the often-static approach of linking processes to outcome in research (Kazdin, 2009). For instance, positive changes were attributed to the possibility of patients to become actively involved in therapy and therapy to be adjusted to patients' needs (i.e., reach a congregation). Based on the findings in this study, we reflected upon the conceptualization of recovery or function and dysfunction in the outcome literature. Instead of a dichotomous understanding, the patients in this study found resignation in a "positive" health definition (Delespaul et al., 2016): health is explained in terms of a person's abilities despite psychological, physical and social challenges. Consequently, the study highlights the importance of an idiosyncratic, individual criterion of evaluation, instead of a normative (external) criterion, something the patients in this study rejected.

The tension between outcome research on the one hand and clinical practice on the other was further addressed in the following chapters, which presented studies conducted in the context of an RCT on the treatment of patients with major depression (Meganck et al., 2017). Here, we focused on the often-voiced demand for standard outcome studies to increase their relevance for clinical practice. Key in this endeavour has been the development of individual-patient indices of clinical significance, with the twofold evaluation described by Jacobson and Truax (1991) being the most widely used method in psychotherapy research (Ogles, 2013). Relying on this influential framework, we expanded standard pre-to-post outcome evaluation with patients' narratives.

In chapter III and *chapter IV* (Study 2), we zoomed in on the positive side of the outcome continuum, examining those patients showing "good" outcome on the BDI-II (Beck et al., 1996; van der Does, 2002) in the RCT. In two separate chapters we scrutinized the experiences of those patients marked as "recovered" (*n*=28) and "improved" (*n*=19) (chapter III) and the underlying change processes as described by those patients (chapter IV). A grounded theory analysis on recovered and improved patients' narratives revealed that patients felt empowered, found personal balance but also encountered ongoing struggle. Experiences of improvement and/or remaining difficulties indicated that from patients' perspective "good' outcome is experienced as an ongoing process and is characterized by a lot of variation in experiences. Remarkably, part of the patients classified as "improved" clearly indicated they did not feel better and similarly for some recovered patients their core problem did not alter. As such, *chapter III* showed that the JT classification does not account for patients more pessimistic perspectives on outcome; especially the heterogeneity amongst the improved sample questioned whether they could be considered under the same "good outcome" header. Consequently, we concluded that such statistical indications of clinical meaningfulness are interpreted warily and ideally contextualized within personal narratives.

In order to gain insight into the processes underlying "good outcome" and the observed differences between recovered and improved patients' experiences, we elucidated the underlying change processes in *chapter IV* as perceived from patients' perspectives. Recovered and improved patients mentioned both helping and hindering influences of therapy, the therapist, patient and external context. As such, "good outcome" did not mean an "all" good process, according to patients' experience of ongoing struggles, and the observed heterogeneity in experiences. Specifically, therapy was experienced as stimulating (i.e., insights, understanding, empowering interaction), relieving (i.e., cathartic effect of talking) but also oftentimes unclear in its influence (i.e., disappointing, uncertain or helping in combination with

other means, e.g., doing sports). Differences between the recovered and improved samples pointed at the importance of patients' agency or active position in the process as well as patients' readiness to change, influenced by external and internal obstacles (especially in the improved sample). Finally, the therapist was attributed an important role and position in the therapy processes as the one who set the right condition. Hence, these findings indicate the variety of co-occurring change processes in and outside of the therapy room and as such a holistic experience of psychotherapy. These findings challenge the idea of a linear and similar trajectory of change and therapy for all patients, pointing at varying needs based on internal and external influences. Consequently, based on the observations regarding the intricate relationship between measured outcome and patients' experienced change processes, we recommend more fine-grained mixed methods studies are required.

In *chapter V* and *chapter VI,* we addressed the conceptual issue regarding the definition of "poor" outcome (Lampropoulos, 2011). Given the lack of a uniform definition of what constitutes negative outcome or negative effects in psychotherapy, the field interprets negative change in terms of a lack of symptom relief and symptomatic worsening (Dimidjian & Hollon, 2010). *Chapter V* presented a mixed methods study on non-improved patients' experiences. Nineteen patients who showed no reliable change on the BDI-II (van der Does, 2002) were selected from the RCT. A grounded theory analysis revealed that "no change" based on outcome measures did not mean a lack of change from patients' perspectives, but rather a partial change or a not enough change. Patients mentioned positive changes on the level of feeling mentally stable or gaining personal strength and insights. On the other hand, patients felt further need to change yet were unable to do so. The core category resulting from patients' narratives in this sample was the feeling of being stuck between knowing versus doing, which resulted in a stalemate. A balance between a "plus" and "minus" was also reflected in the role of therapy, which on the one hand stimulated patients' self-reflection and offered guidance, yet on the other

hit a certain limit or missed something in patients' experience. The therapist was described as predominantly benevolent in the process while patients described their own involvement as characterized by resistance. Similar to the twofold influence of therapy, patients' external context was referred to as an important stimulant to change but also an important impeding stressor. The findings of this study reflect a more nuanced understanding of "no change" which is typically considered a "poor" outcome. Moreover, the heterogeneity in experiences in the samples showed that non-improvement in psychotherapy is more complex and diverse than can be grasped by a lack of symptom reduction.

In *chapter VI*, we provided an in-depth examination of three deteriorated cases, which allowed for the exploration of the relationship between outcome measures and patients' experiences in the case of poor outcome in more detail. Similarities across patients' experiences were found with regard to residual uncontrollable complaints, non-internalized insights and enduring interpersonal difficulties. Interestingly, all three cases had a different outcome and therapy experience, expressing feelings of ambivalence after an unfinished process, disappointment with a lack of change resulting in a premature ending of the process, and a feeling of necessary "deterioration" as a sign of "being right in the middle of the process" and a precondition for structural change. Accordingly, differences were also observed with regard to the influencing factors that were mentioned by patients. Specifically, talking in therapy was experienced as heavy digging or repetitive talking, in general, therapy has not (yet) been able to provide sufficient answers and the therapist was experienced as an ally, silent listener or "too kind". Across the cases, the presence of interpersonal difficulties was remarkable, as well as the – seemingly consequential - tendency to take a distance from others (wanting to be left alone). Patients' agency (broader involvement) in the process as well as extra-therapeutic demands seemed important to understand why therapy was continued or ended. Concluding, this study thus stressed that outcome scores should be interpreted within the narrative of patients

in order to allow for adequate interpretations of what an increase in self-reported symptom severity implies with regard to patients' functioning and the therapy process.

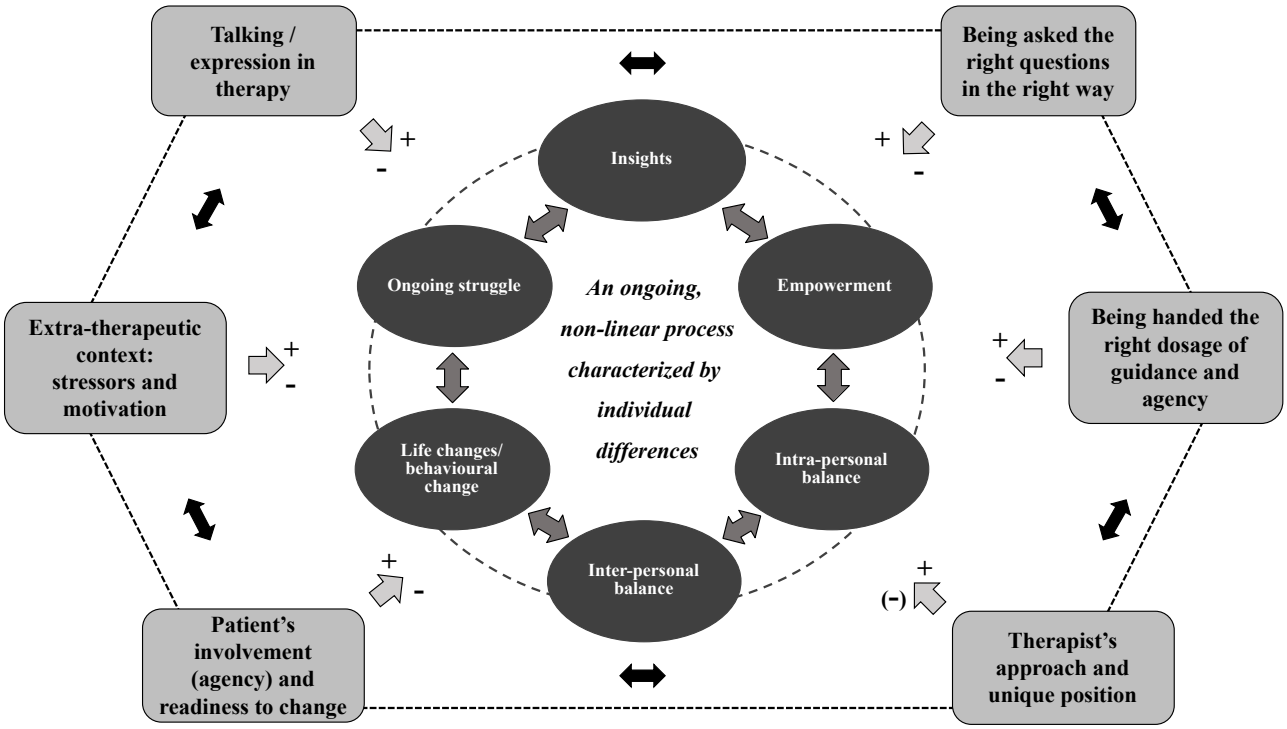
Taken together, by conducting four mixed methods studies on patients' perspectives, this dissertation attempted to open the "black box" of psychotherapy outcome and examined what lies within and beyond the dominant statistical classification of outcome in terms of clinical significance. What we observed across all studies is a more nuanced and diverse picture of outcome and change processes painted by our patients in comparison to the seemingly straightforward statistical classifications. In what follows, we will integrate the key findings from the separate studies, answering the central research objectives: (1) how outcome can be understood from patients' perspectives, (2) which change processes are underlying according to their view and (3) what this implies for the statistical outcome approach and its ability to be useful to inform actual clinical practice. Hence, we resume that our aim is twofold: on the one hand we aim to contribute to a deeper understanding of outcome and the potential underlying change processes (objective 1 and 2; clinical/conceptual), which, in a second step, allows us to reflect on the implications for dominant research practices (objective 3; methodological).

Integrative conceptual model of psychotherapy outcome: Synthesis of patients' perspectives

A critical integration of the findings of the separate studies is aimed at answering the following specific research questions that were addressed throughout the different studies: 1) how can outcome be understood from patients' perspectives? 2) which changes are experienced by patients? and 3) which processes influenced changes according to patients? Our findings show that from patients' perspectives, outcome is an *ongoing, non-linear process characterized by individual differences*, comprising a *holistic experience of multi-dimensional change*. Moreover, *multiple change processes – inside and outside therapy – interact over the course of time* and hence, psychotherapy *outcome* does not necessarily reflect experienced psychotherapy

effects. These findings are summarized in an integrative conceptual model (see Figure D) and discussed consecutively in the following sections.

Figure D. Integrative conceptual model grounded in patients' perspectives



"Outcome" as an ongoing, non-linear process characterized by individual differences

First and foremost, in all studies (in the RCT and naturalistic setting), patients described an *ongoing process* rather than having reached an "end point"; none of the patients with successful outcome results stated they felt "cured", for instance. Rather, patients described remaining difficulties in terms of a process that would continue with or without further (required) help from therapy. Hence, in patients' understanding "outcome" (as evaluated pre-to-post therapy) seems a matter of "the punctuation in the sequence of events" (Watzlawick, Beavin, & Jackson, 1967 p. 94) as highlighted by Greenberg and Pinsof (1986) in reference to Kiesler (1971; 1973) who recognized the arbitrariness of any chosen point to measure outcome. Across the investigated samples, patients moreover often expressed a *non-linear* pattern, for instance characterized by ups and downs, fluctuations, relapse or even initial deterioration. By

this, they moreover expressed important *individual differences*: in each study we extracted different perspectives within the samples with regard to whether patients themselves felt improved or not (broader: how they experienced their outcome at therapy ending). The examination of change processes, moreover, identified important patient, context and patient-therapy/therapist characteristics that indeed support the observation that there is not one single or linear trajectory towards a similar outcome (see Evans & Fletcher, 2013; Kazdin, 2009).

"Outcome" as a holistic experience of multi-dimensional change

Concretely, across samples we observed that patients' experienced changes did not comprise a single change but ranged across multiple facets of patients' self, life and relationships with others, which is in line with previous qualitative outcome studies (e.g., Binder et al., 2010; McLeod, 2013). Moreover, these changes were experienced as interrelated, rather than occurring in isolation (indicated by the dotted circle in the centre of the conceptual model in figure D). Hence, in accordance to what was concluded in the meta-analysis by Levitt et al. (2016), patients described a *holistic* lived experience consisting of co-occurring changes.

Insights, Empowerment, Intra-personal balance, Inter-personal balance, Life changes/behavioural change and Ongoing struggle. Specifically, across studies and samples, patients mentioned the following domains of change: insights, empowerment, intra-personal balance (e.g., relief-experiences), interpersonal balance (e.g., harmony), life changes/ behavioural change (e.g., skills to deal with negative thoughts) and ongoing struggle (i.e., a lack of change or enduring difficulties). We will describe these dimensions hereafter, discussing both similarities and differences among the patients in the different samples.

Insights. Gaining insights or self-understanding is one of the most established findings across studies, both in patients' experiences (Timulak, 2007) and the broader (theoretical) literature (Castonguay & Hill, 2012), and is considered an important mechanism in obtaining positive outcome (Jennissen et al., 2018). Generally, it is believed that insights can occur

suddenly, accompanied by a sensation of discovery, or an "aha" moment as well as develop slowly over time (see Carey et al., 2007). In the present dissertation, this sudden "aha"-experience was particularly described by patients in the inpatient setting who experienced what they called "a revelation". However, at the same time, they also described a slow process of realization in the years following the intense inpatient treatment. In general, the patients across all studies, and in nearly all samples, referred to insights in terms of an increased understanding of themselves, their difficulties and their past. Differences between patients, however, point at the level of *internalization* of insights. The deteriorated cases, for instance, described insights in terms of "interpretations by the therapist" or "feedback", and similarly, some improved patients described insights in terms of an "echo of the therapist". Moreover, while insight is typically considered a key aspect of psychodynamic treatment (Barber, Muran, McCarthy, & Keefe, 2013), in our study, insights were mentioned both by CBT and PDT patients (in line with Carey et al., 2007). Differences regarding the nature of these insights mainly suggest self-understanding and connection between past and present events were typical in PDT and understanding one's problems (e.g., schematic overview) was characteristic for CBT (see Nilsson et al., 2007 for similar findings).

Empowerment. Gaining self-confidence or personal strength over the course of psychotherapy is a robust finding across qualitative outcome studies (Timulak, 2007) and a key observation across various settings and contexts (cf. Delgado-Romero et al., 2016). In the present dissertation, patients from the inpatient setting expressed an "altered self", as they truly felt like "a different person", more independent and able to stand up for themselves (and this in varying degree until five to six years after ending therapy). Similarly, patients who recovered and improved in symptom severity in the randomized controlled study, described increased self-confidence, gained coping skills and emancipation. This was most typical for recovered patients and appeared still in development or less established for part of the patients marked as

"improved". Likewise, in the sample of non-improved patients, a growth in mental stability and personal strengths were mentioned. However, both patients who were marked as "improved" and "non-improved" also experienced feelings of "inability" and doubt (cf. life change/behaviour changes). Particularly for the cases who deteriorated in symptom severity, "empowerment" seemed absent or limited as they especially experienced remaining difficulties as "uncontrollable". When comparing the experiences of the patients across all studied samples, differences among patients seem to suggest the role of patients' level of agency, which is often (theoretically) related to empowerment (e.g., McAdams, 2001) and readiness to change (Norcross et al., 2011). This will be further discussed below (cf. patients' involvement).

Intra- and inter-personal balance. In all but one sample, we differentiated both intra-personal and inter-personal changes. Reaching intrapersonal balance was expressed in terms of feeling relieved, becoming milder towards oneself and generally feeling calmer (in line with findings synthesized by Timulak, 2007), which was most typical for recovered and improved patients; the inpatients described themselves as "mollified" (cf. feeling like a "different person"). Consequently, intrapersonal changes seemed to result in or coincide with improved interpersonal relationship. Typical for both recovered and improved patients was the experience of a better harmony and positive relationships/interactions, although this implied a separation for many of the improved patients. Patients in the nonimproved sample remarkably did not typically mention interpersonal changes, however, and deteriorated cases experienced increasing interpersonal difficulties. These differences could point at the importance of improvement on a symptomatic level to enable changes on an interpersonal level. However, it might as well suggest the role of interpersonal difficulties and their impact on depressive symptoms and broader intrapersonal unbalance. For instance, Mohr (1995) found that particularly patients who exhibit high levels of interpersonal disturbance are at risk for deterioration in therapy. On the other hand, interpersonal difficulties are considered to be at the

heart of depression experiences (Joiner, 2000), and thus, the direction of the relationship seems unclear, and is thus plausibly more complex (cf. infra). Therefore, in figure D we present and assume a mutual influence corresponding to patients' holistic experience of co-occurring changes.

Life changes/behavioural changes. In all our studies, patients mentioned changes that concerned their lives, which reflected actual "life changes" (e.g., choosing for a new career path) and certain behavioural changes (e.g., dealing things differently, cf. coping skills). Similar to previous observations (Nilsson et al., 2007) CBT patients in the recovered and improved sample described to have learned to cope with difficulties during the directive treatment (Nilsson et al., 2007). These tools especially concerned perceiving and handling (negative) thoughts and emotions differently. However, life changes were also observed on the level of becoming more self-centred (and thus bothering less with the troubles in someone else's life), which was experienced as a "healthy necessity" (both mentioned in CBT and PDT). Differences among patients (across samples) seemingly occur with regard to patients' ability of translating these insights and tools from therapy into actual life changes. For instance, patients in the non-improved sample typically stated the suggestions from therapy often hit a limit (e.g., impossible to use certain tips when feeling bad). For the deteriorated cases, it was particularly remarkable how insights or tips coming from the therapist were deemed "irrelevant" or even "incorrect" by one patient (i.e., Miriam, see chapter VI) who experienced therapy as insufficient. Life-altering changes were particularly observed for patients in the naturalistic study who engaged in one-year of intense inpatient treatment, indicating the difference between changes obtained during time-limited outpatient treatment and long-term inpatient care (cf. influence of the level of intensity of treatment; see Cuijpers et al., 2013 for a discussion) A thorough discussion on the unique case of inpatient care as well as group psychotherapy is beyond the scope of this dissertation, however, we refer the reader to Burlingame, Strauss and Joyce (2013).

Ongoing struggle. Strikingly, across all outcome studies and outcome categories (from recovery to deterioration), remaining difficulties were mentioned. This was most typically for deteriorated cases and literally described as a feeling of being "unable" to make further change by several non-improved and improved patients. Nonetheless, also recovered patients typically stated they were "not there yet", as the process of change was still ongoing and/or residual complaints were experienced. Comparing the findings across patients and samples highlights patients' perspective regarding these remaining difficulties differed. For instance, while some patients (typically in the recovered sample and inpatient group) could consider remaining issues as "normal", other patients seemed preoccupied with remaining difficulties, sometimes impeding the therapeutic process (typical for patients marked as improved, deteriorated cases and also mentioned by non-improved patients; see external stressors in the extra-therapeutic context, cf. infra). However, importantly, the ongoing struggles experienced across all samples and outcome classifications did not correspond to patients' further need for therapy in a one-on-one relationship, that is, ongoing difficulties did not necessarily mean patients wanted further treatment. We will return to this observation later on in this discussion.

In summary, differences among the patients studied in this dissertation were observed with regard to the level of empowerment, internalization of insights, reaching intra- and/or interpersonal balance, the extent to which therapeutic tools and insight could be translated into actual life or behavioural changes and whether ongoing struggles were perceived as normal or impeding and required further treatment or not. In comparing patients' experiences across the samples, we have attempted to give voice to both the *within-group differences* observed in each sample and the remarkable *similarities across* the samples. We refer to the useful framework proposed by Hill et al. (2005) and criteria for larger samples by Knox et al. (2006) to describe the *salience* of experiences within samples based on a qualitative methodology and semi-structured interviews (see chapter III and IV). We will return to these observations when

discussing the implications of our findings for the classification of outcome and whether or not this allow for meaningful differentiation of patients' functioning post-therapy.

Multiple change processes – in and outside therapy – interact over the course of time.

Based on patients' narratives, differences in experienced changes particularly point at the role of patients' level of agency (or internal obstacles), the interaction with the external context and the extent to which therapy/therapist and patient were able to congregate over the course of treatment. Patients across all studied samples moreover indicated that not one factor could explain the experienced changes. This is depicted in Figure D by the dotted line connecting the influencing factors and the multiple arrows indicating their influence and interrelatedness. Moreover, both positive (+) and negative (-) influences were mentioned for all influencing factors, in correspondence to the mixed experiences of change across all studied samples (see Figure D).

The role of therapy: being asked the right questions in the right way, talking and expression, being handed the right dosage of guidance and agency. Across all studies, patients addressed the role of therapy in three general ways. First, in all studies, therapy was described as stimulating patients' thinking, which consequently lead to new insights and self-understanding. In the meta-analysis by Levitt et al. (2016) similar findings were summarized as "reflexivity" leading to "holistic awareness and a new self-narrative, abetted by therapists' insights" (p. 819). Here, they point at the active role of the therapist helping to gain new insights, which corresponds to the narratives in our studies: gaining insights was attributed to the simulating role of the therapy/therapist. Specifically, comparing the patients across the samples, suggests the importance of *being asked the right questions in the right way.* For many patients who were marked as recovered, therapy offered a confronting mirror or helped to see connections and overview. For two deteriorated cases, the exploration in therapy felt more like

"digging" (cf. supra), while in the non-improved sample, several patients mentioned the role of reflection, being provided with insight and being asked the right questions by the therapist.

Secondly, *talking or expressing emotions* was considered important, leading to feeling relieved and finding support (e.g., being able to talk; see Timulak, 2007; 2010). Interestingly, while recovered patients typically described this experience as cathartic, deteriorated cases mentioned a "heavy digging" (i.e., talking as a way to bring up the heavy past) which made them feel "worse" (e.g., overwhelming emotions). For some, talking was however mere repetitive (e.g., case Miriam in chapter VI), and dissatisfied patients in CBT generally thought therapy was "too superficial" (e.g., only focussed on present difficulties, not including the broader story; see also Nilsson et al., 2007 in this regard). The patients from the inpatient setting add to this observation, that this relieving role of therapy also counted for *other* forms of expression, particularly creative forms of therapy (e.g., art, dance, movement, music) were experienced as a very powerful means, especially when expressing feelings verbally was difficult. In that sense, our findings support the hypothesis that a "cathartic" experience is important to facilitate further change indicating different levels of "gaining insight", for instance insights as provided by the therapist and "echoed" are qualitative different from those occurring through an "aha" or *lived* experience (Bohart 1993; Castonguay & Hill, 2012).

Third, differences were observed with regard to patients' required need for guidance and autonomy. In that sense, therapy was evaluated in terms of *being handed the right dosage of guidance and agency.* Typically, recovered and improved patients stated that they valued how therapy actively involved them. Particularly in CBT, the interaction with the therapist was valued (i.e., therapist took up an active role, offered direction and feedback), and in both CBT ad PDT the patient felt empowered to take the lead in their own process (cf. the therapist did not instruct on what to do, stimulating patients' *own* responsibility). This need for active participation was perhaps most typically mentioned by patients in the inpatient setting, as the

therapy setting explicitly required patients to take on an active role during their stay (cf. requiring a certain stability and personal strength to take the initiative in group therapy or asking others for help). For some, this offered too little guidance or direction and made therapy difficult (especially for patients identified as "dependent-passive"; cf. patients' involvement). Similarly, dissatisfied patients in PDT who were studied in context of the RCT, typically described to have missed concrete tools, tips and direction during their therapy (see also Nilsson et al., 2007 in this regard).

Patients' involvement (agency) and readiness to change. The above-mentioned experiences of therapy appeared strongly related to patients' own involvement in the therapy process. Patient characteristics are considered the most important predictor of outcome (Bohart & Wade, 2013), particularly suggesting the role of patient's involvement or "agency" (McAdams, 2001)[33] and readiness to change (cf. stages of change models; Norcross et al., 2011). In all samples, we differentiated patients who described themselves as assertive or actively involved and patients who mentioned internal obstacles (most often convictions of impossibility to change and resistance) and passivity. In the inpatient sample this resulted into a differentiation of four patient positions that express these (patient-experienced) degrees of agency: assertive-active, over-active, dependent-passive and independent-keeping a distance. Such characteristic differences were not observed in the samples of the RCT, suggesting differences in populations and therapeutic experiences (cf. inpatient versus time-limited outpatient therapy). In the RCT, both non-improved and improved patients typically described experiencing difficulties (i.e., resistance) in the process pointing at internal obstacles to engage in therapy. In contrast, patients who recovered in symptom severity typically described themselves as actively involved in the therapy and change process. Notably, also if they did not find therapy helpful, they actively engaged in activities that would help them get better,

[33] Cf. "Agency encompasses a wide range of psychological and motivational ideas, including the concepts of strength, power, expansion, mastery, control, dominance, autonomy, separation, and independence." (McAdams, 2001, para 7)

suggesting an already more empowered or agentic position prior to entering therapy (cf. readiness to change; Norcross et al., 2011). Deteriorated cases were characterised by passivity and minimal involvement (e.g., case Mark who seemingly put his experiences at a distance), active involvement (e.g., daring to ask therapist for certain interventions, case Ria) and "disobedient" to the therapy provided and impeded by "impossible solutions" (case Miriam). As most typically explained by patients in the inpatient setting, what seems important is the coming together or congregation of patients' needs/agency (e.g., motivation and expectation) and the therapy offer. Especially the therapist was attributed an important role in this process, as the one who set the right conditions for therapy and change (cf. patient-therapist dyad as a *dynamic* relationship, Clarkin & Levy, 2004). We refer readers to the contextual model as developed by Wampold and Imel (2015). This theoretical model of psychotherapy differentiates three pathways that lead to change and places the relationship between patient and therapist central. In particular, the extent to which patients can move beyond their own explanations for their distress (often hindering their improvement, characterized by an "impossibility") and engage in the rationale offered by therapy, is considered an important prerequisite for the therapeutic process and change to take place.

Therapist's benevolent approach and unique position. Although inevitably related to the therapy as such, in all the conducted studies, the *person* of the therapist was mentioned explicitly. This corresponds to a "common factors" model that sees the therapeutic relationship (and thus the therapist) is one of the key factors in achieving change in therapy, rather than the active ingredients of therapy (see Wampold & Imel, 2015). In literature, this is predominantly referred to as the "working alliance" between therapist and patient, to date considered as one of the stronger predictors of outcome (Bordin, 1979; see however Hatcher & Barends, 2006 and Ross, Polaschek, & Ward, 2008 for a critical discussion on the limitations of current alliance measurement and use of Bordin's definition). According to patients' perspectives in the present

dissertation, a "good therapeutic relationship" was described in terms of a non-judging therapist who allowed patients to open up, an encouraging and acknowledging therapist who empowered patients and the therapist who asked the right questions which stimulating patients' self-reflection. Remarkably, this was mentioned across all samples and thus outcome classifications. When dissatisfaction about the therapist was expressed this seemed individual to the patient and patient-therapist dyad, rather than related to patients' symptomatic outcome. Also, and in line with previous observations (McLeod, 2013), patients rarely overtly expressed negative feelings about the therapist, rather attributing it to the approach than the intentions by the therapists (although our analyses did not allow to study these experiences in-depth; we refer the reader to literature on alliance ruptures, Safran, Muran, & Eubanks-Carter, 2011). Differences between CBT and PDT therapists were however mentioned: CBT therapists were often perceived as the active participant and initiator of the sessions while PDT therapists were seen as stimulating an ongoing mental process on the side of the patient in a more subtle way (see also Nilsson et al., 2007).

The extra-therapeutic context: stressors and stimulant. Finally, in every study of this dissertation, patients referred to the impact of their personal context on their change and therapy process. The "extra-therapeutic influence" is one of four factors included in the "four-factor model" by Lambert (1992; besides "common factors", "technique factors", and "expectancy factors"), and it is estimated to contribute 40% to change (Miller, Duncan & Hubble, 1997). In the present dissertation, especially the interpersonal (significant others) and professional context were deemed both a source of meaning and support as well as the reason for (ongoing) difficulties (cf. Drisko, 2004). This observation is important as it broadens the understanding of change and change processes from the in-session setting (and thus the role of the therapist, alliance and therapeutic techniques or patient) to the context *outside* of treatment. As found by Gianakis and Carey (2008) and observed by Slade and Longden (2015), people suffering from

mental difficulties can (and often) change *without* therapy, and the effects of therapy – "that one hour a week" (as one of the patients in our study remarked) – are perhaps indeed more modest than sometimes assumed, especially in causal research designs like RCTs (see Evans & Fletcher, 2013). As such, our findings show that, from patients' perspectives, the therapy process is characterized by a holistic experience and perceived in light of and interaction with patients' broader life context. Supporting the observation that the comparability of patients' change is not straight-forward as different external and internal influences define different patient needs and thus therapy process.

Summary. We thus conclude that psychotherapy, according to patients' experiences, can be understood as a holistic experience – correspondingly patients' holistic multi-dimensional experience of change (cf. Levitt et al., 2016) – as it is embedded in patients' life, person and external context. Understanding the effects of psychotherapy hence requires contextualization within co-influencing factors. As such, our findings suggest an intricate relationship between patients' observed symptom severity, patients' experiences of change and the role of therapy and warns for simple process-outcome conclusions. Answering the question whether therapy was successful for a certain patient and why, hence becomes a balanced one, requiring that we take into account multiple factors and circumstances. Clearly, agency will not always lead to good outcome, like a good therapeutic relationship does not – on its own – guarantee successful outcome. This consequently challenges the linear and static evaluation in process-outcome research (Kazdin, 2009) and calls for a contextualized evaluation and understanding of therapy effects. Especially this kind of contextual information is omitted in RCTs (that aim at describing the sample in terms of similarities, e.g., "depression", and the setting of therapy based on its standardized format, e.g., "the protocol"), which are therefore strongly criticized with regard to their clinical relevance. Hence, we have illustrated the clinical value of looking at these seemingly homogeneous samples and standardized contexts *within* a

highly standardized and controlled setting from an idiosyncratic and contextualized perspective (see also Dattilio et al., 2010). We will return to this when discussing some of the avenues we believe are promising for future research.

Psychotherapy *outcome* does not necessarily reflect experienced psychotherapy *effects*

As the above-mentioned influencing factors indicate, patients' outcome did not necessarily reflect the experienced *therapy effects*, meaning that the measured pre-to-post outcome did not necessarily, or exclusively, reflect the effects of therapy according to patients' perspectives. This poses a challenge to the frequent use of "outcome" and "therapy effect" as synonyms (Carey et al., 2006; Lampropoulos, 2011). In all of our studies, the patients' perspectives nuanced the role of therapy, by describing what we called a more holistic experience of therapy (e.g., interacting with patients' extra-therapeutic world) or by expressing their (dis)satisfaction about the effects of therapy (e.g., recovered patients who thought therapy was insufficient and attributed change to other factors). Nonetheless, in the outcome literature, measured outcome and psychotherapy effects are often used interchangeably, while outcome is not necessarily the result of treatment alone (cf. Lambert et al., 1992; four factors model). Carey et al. (2006) therefore argue that, in order to grasp the phenomenon of change (or "outcome") and the effects of therapy, it is warranted to understand them as related *but* distinct phenomena and study them accordingly (Carey et al., 2006). Doing so they highlight the (impact of a) lack of a clear definition of "change" to date; we will return to this conceptual lacuna in the field later in this discussion.

The clinical (un)meaningfulness of statistical outcome classification

We now turn to the question whether the influential statistical framework of clinical significance by Jacobson and Truax (1991) provides a clinically useful indication of psychotherapy outcome in light of patients' own experiences (cf. Atkins et al., 2005; Kazdin, 2001a). As we outlined in the general introduction of this dissertation, the definition of clinical

significance was developed with the aim to find "agreed-upon conventions for determining what constitutes improvement in an individual, and, if possible, general consensus as to what is meant by clinical significance." (Jacobson et al., 1984, p. 338-339). Answering the question of what can be considered "clinically significant" inherently depends on how outcome is defined. To date, there is no uniform definition of what desired outcome in psychotherapy is, so it typically varies depending on the theoretical orientation and inherently contains a normative judgement that is perspective-dependent (Strupp & Hadley, 1977). In psychotherapy research, this lack of conceptualization has resulted in a strong reliance on statistical convention and objective measurement (see Malterud, 2001; McClimans, 2013; Greenberg & Pinsof, 1986) to which the JT classification offered an important contribution in the search of finding a widely applicable approach (Jacobson et al., 1984). We resume that the JT classification departs from the understanding of psychotherapy outcome as a continuum yet (dichotomously) divides patients into a functional and dysfunctional population with regard to a certain target problem evaluated by means of a specified instrument. In the present dissertation, we followed the dominant primary outcome method for evaluating outcome (De Los Reyes et al., 2011) and, given the focus on major depression in the RCT, the BDI was used as a psychometrically sound measure of depression severity to operationalize the clinical relevance of the measured individual-patient change according to the JT classifications.

Whether the JT method provides a clinically meaningful indication of psychotherapy outcome in light of patients' own experiences can be brought down to the following two questions: 1) does it grasp meaningful change in light of patients' experiences (i.e., daily life and therapy; cf. Atkins et al., 2005; Kazdin, 2001a) and 2) does it permit a clinically useful differentiation of patients' functioning after therapy? (cf. McElvaney & Timulak, 2013)

"Clinical significance" versus the multiplicity and nuance in patients' experiences

With regard to the first question, our findings showed that patients experienced outcome as an ongoing, fluctuating and non-linear process, comprising a holistic experience of multi-dimensional change. In correspondence to previous observations, a pre-to-post, one-dimensional evaluation of outcome underlying the JT method falls short to grasp this multidimensionality within patients' experience (Hill et al., 2013; McLeod, 2013). Doing so, our studies confirm previous observations that symptomatic changes only comprise a small proportion in patients' own narratives (Levitt et al., 2005; McLeod, 2013) and as such, the reduction of symptoms only proofs to be a poor indication of successful outcome in light of patients' perspectives (Zimmerman et al., 2012). Conversely, an increase in scores is not necessarily deemed negative by patients (as expressed by the deteriorated case Ria described in chapter VI). However, our findings do not just illustrate that patients' experiences are more complex than can be grasped by quantitative outcome measurement (as outlined in the general introduction, our aim is to focus on exploring *what* can and cannot be grasped by means of the JT classification and whether and to which extent this does allows for clinically meaningful conclusions). Rather, our findings challenge the tenability of the categorizations as such, as in patients' experiences, outcome was not as clear cut, "good" or "poor" as suggested by the outcome categories. Patients' experiences were both negative and positive across all samples and outcome classifications, hence warning us for the simple and unnuanced use of terms like "good" and "poor" outcome, "recovery", "improvement", "no change" and "deterioration", as oftentimes they did not correspond to patients own more *ambiguous outcome experience*.

From patients' perspectives we can thus wonder whether it is worthwhile to classify them into mutually exclusive categories at all. As the patients in our studies show, outcome classification based on solely one criterion (e.g., pre-post symptom reduction) clearly falls short to grasp the multiple facets of outcome. When evaluating patients on multiple dimensions (e.g.,

insight, empowerment/agency, intra-personal and inter-personal balance, etc.), we can expect that a patient's outcome classification will *vary* depending on the dimension he or she is being evaluated on (e.g., a patient can score high on gained insight, but low on empowerment

Evaluating the clinical meaningfulness of the JT classification also requires to consider the impact of the utilized measure in our study (i.e., BDI), however. We can indeed wonder whether the use of a different instrument would have led to different conclusions (Atkins et al., 2005)[34]. Indeed, the arbitrariness of outcome measures is a major dispute in the field (Kazdin, 2006), and the use of multiple measures is increasingly advocated to tackle the limitations of using a single instrument (Hill et al., 2013). However, the interpretation of possible discrepant findings on different outcome measures still confuses the field (see De Los Reyes, 2011) and puzzles the decisions regarding clinical significance (Atkins et al., 2005; Ogles, 2013). Hence, it seems that, in order for outcome findings to be clinically meaningful, it is of essential to scrutinize the reasons *why* outcome findings point to certain, perhaps different, directions. A multi-dimensional understanding and evaluation of outcome as such would imply an interpretation of the multi-facetted nature of outcome rather than reduction to one primary outcome measure.

"Clinical significance" versus cross-sample similarities and within-sample differences

With regard to the second question, whether JT allows for meaningful differentiation of patients' functioning pre-to-post therapy, patients' experiences in our studies revealed that important *cross-sample similarities* as well as *within-sample differences* were remarkable. Firstly, in line with previous observations, the similarities among the samples seemed to outweigh the differences (see McElvaney & Timulak, 2013). In comparing patients'

[34] Note that Jacobson et al. (1984) themselves indicated that the definition is as "strong" as the measure that is being used. However, their definition of "strong" relies on the internal consistency of the measure imbedded in the calculation of the RCI. Based on those criteria, the BDI would be regarded a decent measure, and hence is therefore often used. However, the argument being made here addresses the strength of the measure in terms of its level of arbitrariness to evaluate outcome in light of patients' experience (cf. Blanton & Jaccard, 2006; Kazdin,2006).

experiences, the clearest distinction could be observed between the sample of recovered patients (in which the majority of patients stated that they could go on without further help, see chapter III and IV) and the patients who (clearly) deteriorated (see chapter VI). In light of a first-person patient perspective, our results thus indicate that the functional-dysfunctional distribution-approach (based on self-reported symptom severity) showed most useful to indicate general differences with regard to the patients at the extreme ends of the outcome continuum. Patients who clearly improved or "recovered" in terms of symptom severity and those who clearly worsened in symptom severity (which came down to one patient, Ria cf. chapter VI) most clearly corresponded to the classification they were allocated to. However, also for those patients, the classifications "recovery" and "deterioration" did not necessarily correspond to their *own* valuation of outcome (e.g., Ria felt "right in the middle of the process" and thus "on track" rather than alarmingly "off track", while several recovered patients did not think their core problems were altered). The differences between the recovered and improved sample and the improved and non-improved sample were particularly less clear cut. Specifically, the JT classification could not account for the more pessimistic experiences observed among those patients marked as "improved" and missed the positive changes for those defined as "unchanged". Especially large heterogeneity amongst these samples problematized the classification of those patients in one and the same group (especially in the improved sample). Based on patients' perspectives, the separation of "good" and "poor" outcome hence becomes fuzzy and mainly points at the heterogeneity *within* the different samples.

In statistical terms this suggests that *reliable change* (i.e., RCI) is not sufficient to differentiate patients' *functioning* in a meaningful way. For instance, in our sample, reliable deterioration seemed only in line with the subjective experience of deterioration in the case who nearly doubled in symptom scores (i.e., Ria, see chapter VI). Related to this is the critique on the JT method that its categories are unreliable at the category ends, making the classification

of samples more arbitrary than intended. Also, as the JT method does not correct for "regression to the mean" it can be that high scoring patients (like observed for the improved sample) are more likely to show a decrease in scores, regardless of the actual meaning of that change (Lambert et al., 2008). A suggested alternative that addresses this weakness (i.e., that the RCI is not adjusted to a patient's pre-treatment symptom severity) is "percentage improvement" (PI). Instead of passing the 11.3 cut-off (as predefined for the Dutch BDI; van der Does, 2002), this means every patient must improve 50% in comparison to their initial level (Ogles, 2013). While this indeed increases the idiosyncrasy of the level of comparison, it however still relies strongly on the outcome measure and thus brings us back to first problem with the JT method (and outcome measurement in general; Ogles, 2013).

This thus suggests the importance of the second criterion to define clinical significance, namely *normative comparison* (Jacobson et al., 1984): a patient's outcome score is compared to a patient and non-patient normative sample. With regard to deterioration, Jacobson and colleagues (1984) acknowledged that they could not define a clear counterpart to the distributional cut-off for clinical significance, and they encouraged future investigators to suggest some additional criteria to establish reasonable indices of clinically significant deterioration. One such example can be found in the paper by Wise (2004) who suggested more stringent criteria in the case of deterioration. However, listening to patients' voices in our studies, in all categories – from recovered to deteriorated cases – a significant number of patients did not agree to the classification he or she was being attributed to; in other words: *patients did not feel in line with the norm they were being compared to*. This was literally voiced by the inpatient sample: patients found resignation in comparing themselves to an *internal* norm, rather than adopting the standards of (the harsh) society.

Finally, our observations that the outcome classifications become "fuzzy" when comparing to patients' experiences across the samples, reconfirm the typically-voiced

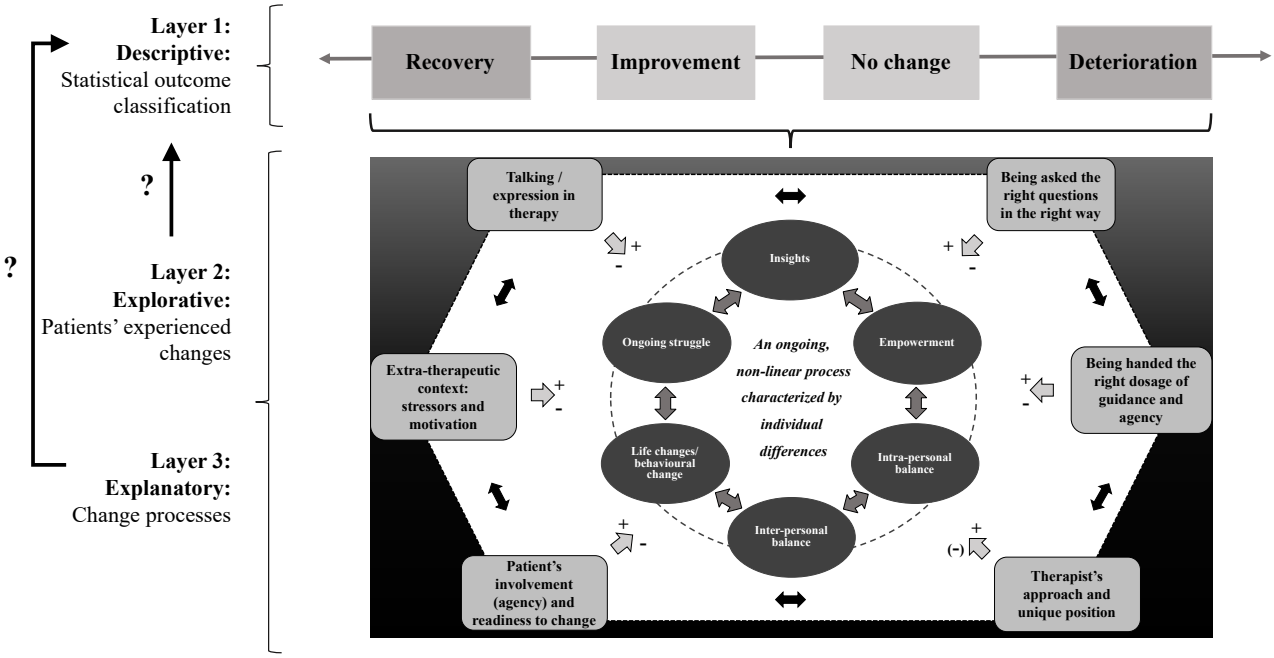
limitation of *between*-group comparison as being only minimally informative about the variation *within* the group (cf. general introduction of this dissertation). As the JT method has explicitly been developed to offer insight into the within-sample variation, in a way they offered an answer to mere mean-level and between-group knowledge. However, when indices of clinical significance are reported in outcome studies and particularly when synthesized in meta-analyses (although still a scarcely applied practice), patients are *grouped* into categories based on their outcome classification and presented in percentages of, for instance, recovered and improved subjects. Based on the cross-sample similarities and within-sample variation observed in this dissertation, we strongly question the *clinical value* of this approach to clinically relevant outcome reporting (cf. Fisher et al., 2018; Westen et al., 2004). As, in fact, the step of *re-grouping* patients seemingly contrasts the aim to provide an *individual-patient* indication of the observed change as intended in the first place (cf. Jacobson et al., 1984).

Ergo, there's no going back after opening the "black box"

Based on our findings, we thus conclude that the influential JT framework of clinical significance falls short of grasping clinically meaningful change with regard to patients' perspectives. Moreover, cross-sample similarities as well as within-sample difference challenge the tenability of relying on strict outcome categorizations as a means to draw clinically useful conclusions. This problematizes the pre-post, pre-defined one-dimensional evaluation of outcome as well as the comparison of patients' outcome to (one and the same) external norm. Overall, differentiating patients' functioning into functional versus dysfunctional distributions falls short of grasping patients' outcome experience and does not allow relevant comparison across patients (van Os et al., 2019). From patients' perspectives, outcome must be understood as a holistic experience of multi-dimensional change, characterized as an ongoing, non-linear process which moreover revealed important individual differences. These differences particularly convey patients' level of agency (or internal obstacles), extra-therapeutic factors

and the patient-therapy/therapist dyad which evolves over time. These conclusions are depicted in figure E, which shows the inverted three-layered model we put forward to study psychotherapy outcome (cf. figure A), concluding that opening the black box of psychotherapy showed to be easier than closing it again. Hence, while the present dissertation attempted to open the black box, it arrives at a conclusion that emphasises the need to continue the search for clinical meaningful outcome interpretations rather than reside in the continued use of "the best available" statistical answer so far (which is mostly motivated by practical consideration; Lambert et al., 2008; Ogles, 2013). Thus, the search for clinical meaningfulness in outcome research continues (see Westen et al., 2004), and below we will discuss some of the avenues we believe are note-worthy and promising.

Figure E. Opening the black box of psychotherapy research (is easier than closing it again)



Note. The image illustrates that when departing from patients' perspectives it becomes unclear how to categorize patients' outcome in the pre-existing classifications, given the multi-dimensional, non-linear and holistic nature of their experiences. More importantly, important similarities across samples and variety within question if it is meaningful to classify patients at all (indicated by question marks).

Can a statistical outcome continuum provide a valuable answer?

An often-proposed solution to the shortcomings of reliance on a distinctive categorization of functional and dysfunctional distributions (like the JT method) worth mentioning in light of our observations, is the presentation of complaints on a single outcome continuum. For instance, authors like Tingey et al. (1996) proposed the use of multiple cut-offs and corresponding normative samples (cf. outcome ranging from asymptomatic to severely symptomatic). Indeed, our findings suggest that working with groups or categories does no justice to patients' varying experiences, as groups unjustly suggest homogeneity among the patients within the groups and suggest the evolution towards some sort of stable outcome (the functional distribution). Whereas patients' in our study experienced ambiguity, change in one domain but not in the other, ongoing struggles, etcetera. While a continuum would solve the first problem – which is the unjust suggestion of homogeneity by 'groups' – it does not solve the second problem of unjustified unidimensionality.

Should we consult the perspective of significant others and bring back "social validity"?

A second approach that is valuable to mention in light of our findings, is the interpretation of clinical relevance in terms of "social validity", which originated (and remains mainly used) in the field of behavioural psychology. Seemingly simultaneous to the statistical developments to clinical significance, a method was developed which relied on the subjective judgement by significant others outside of the therapy or research setting (see Ogles, 2013). Kazdin (1977), for instance, combined both normative comparison (i.e., comparing a patient's functioning to representative peers, similar to the second criterion in the TJ method) and the subjective evaluation by significant others (those who most likely have contact with the patient) or experts.

While we believe the idea of consulting significant others can indeed enable one to get closer to an understanding of patients' change in a real-life manner and hence increase the

clinically meaningfulness, it cannot replace patients' own subjective experiences (cf. holistic, non-linear, multi-dimensional). Moreover, a number of methodological issues impede the usefulness of this method, particularly how social validity is defined and measured, which may explain why this method did not lead to a wide application and influence in outcome research as did the easy-computable JT method (see Ogles, 2013). Most importantly, the concept "social validity" has been used in a myriad of ways leading to a disparate field. Apart from that, researchers have relied on the use of video tapes to rate patients pre-post treatment which, when sufficient representative material is shown, becomes a highly demanding task for external raters (see Foster & Mash, 1999 for a discussion). Lastly, questionnaire-based assessment of significant others' judgement is susceptible to the previously-outlined limitations of outcome measurement. A thorough discussion on this line of work is beyond the scope of this dissertation, yet it is worth mentioning, as it does show that, from the beginning of the search for indicators of clinical relevance, researchers sought for ecologically valid methods apart from the statistical answers. As concluded by Ogles (2013), we agree that promising avenues could originate from this approach to evaluate clinical relevance. For instance, it could provide a framework to evaluate the societal perspective included in the "tripartite model" suggested by Strupp and Hadley (1977), which encourages to study of outcome by examining the perspective of the patient, therapist and society (including significant others).

The remaining need for a workable outcome definition and risk for reification

Hence, although both the idea of an outcome continuum and social validity by relying on the perspectives of significant others seem to bring us closer to the meaning of outcome as expressed by patients in this dissertation, they do not resolve the limitations of a currently lacking *workable* definition of outcome in the field of psychotherapy research (Kiesler, 1966; Strupp & Hadley, 1977). The discrepancies found between patients' perspectives and the statistical classifications suggested by the JT method, in that sense warn for the implications of

taking statistical conventions (as the identification of an individual-patient's outcome) at face value. Specifically, our findings show that this potentially masks clinically relevant nuances and experiences, which entails the risk of providing us with a false idea of understanding patients' functioning after therapy. This risk has been described as methodological "reification" and is criticized for its impact on research findings and the constructed scientific knowledge (cf. Hesse-Biber, 2015). For instance, when the multifaceted nature of outcome is artificially simplified and brought down to one primary outcome measure, the risk is to start understanding the phenomenon "outcome" as a "simple construct" accordingly (e.g., mere symptom reduction), while this may not approximate the phenomenon or object of study we are in fact interested in (e.g., major depression). Major Depression in this case can serve as a topical example. Several critics have warned for the consequences of reducing the understanding of depression as well as the research measures to the descriptive criteria listed in diagnostical manuals (like the DSM; APA, 2013; e.g., Dehue, 2014; Hyman, 2010), as it installs a self-perpetuating process, which *predetermines* our understanding of depression and findings from research.

Thus, the observation (in previous findings and our own studies) that outcome in therapy for major depression goes far beyond mere symptomatic changes, is actually not that surprising in light of the broader discussion on how major depression should be conceptualized to approach the lived-experiences of patients. Based on those lines of research that approach depression from a phenomenological perspective (see Ratcliffe, 2014), it shows that the experience of depression is multi-facetted (e.g., comprising experiences of self, other, time etc.) and moreover characterized by important idiosyncratic experiences (see also De Smet et al., unpublished manuscript; see also Gupta, 2007 on the heterogeneity amongst patients diagnosed with major depression). Consequently, if research into psychotherapy outcome for the treatment of depression would include a *broader* or more *idiosyncratic* understanding of the phenomenon

"depression" to begin with, this would consequently imply a different approach to the conceptualization and evaluation of "outcome" at the end of treatment (cf. De Smet et al., unpublished manuscript). In other words, addressing the issue of finding a workable definition of outcome should concern the broader question: *what (or who) are we measuring outcome for*. We will return to this remark when discussing the implications for research and future avenues.

Towards a "first-person" understanding of clinically significant change

The JT classification offers an individual-patient outcome evaluation by departing from *one and the same* pre-defined norm to judge the relevance of patients' outcome (Jacobson et al., 1984). Our findings based on patients' experiences show that, defining clinical meaningful change from a *first*-person perspective requires an understanding of clinical meaningfulness of outcome based on an internal, subjective rather than an external, "objective" norm. This *first-person* definition of clinical meaningful change comprises the core of a growing emphasis on redefining "clinical recovery" (e.g., the JT method) into "personal recovery". This is summarized by Slade and Longden (2015) as "a process or a continuum; subjectively defined by the person themselves; 'rated' by the person experiencing the mental health difficulties [as the individual is] considered the expert on their recovery [and] recovery means different things to different people, although there are aspects that many people share." (p. 3). This definition resembles our findings that showed individual differences, idiosyncratic valuations of outcome and at the same time important shared cross-sample similarities regarding the domains of experienced change (see also Levitt et al., 2016; see Timulak, 2007; 2010). Moreover, this patient-centred definition moreover presumes that change is in essence an *internal,* private event experienced by the individual patient (see also Carey et al., 2006; Tryon, 2005), which implies that it is not necessarily objectively (i.e., externally) observable, and thus also not necessarily measurable in the same way for all patients (Slade & Longden, 2015). The patients in our study, for instance, emphasized the importance of gaining insight, which did not

necessarily coincide with change at symptom level. Conversely, while some patients felt "symptom free" they did not feel their core difficulties had altered, hence not feeling personally recovered. Empowerment and agency, both as a therapeutic tool and as an outcome in therapy, seem key in patients' own experience (see also Timulak, 2007; 2010). Thus, rather than evaluating whether patients have crossed a predefined cut-off, it seems more relevant to focus on whether patients actually feel empowered or agentic to continue on their own (with or without remaining symptoms). Interesting recommendations for research and clinical practice based on a "personal recovery" idea can be found in the work by Slade (2009).

Related to concept of "personal recovery", a changing definition of health and illness towards an idea of "positive health" can be observed in the health care field, as we discussed in chapter II of this dissertation. Increasingly, the classical health-illness dichotomy (cf. WHO definition of health since 1948[35]) is being replaced by a focus on strengths and abilities, *despite* difficulties, which is considered more in line with many of the chronic difficulties in our society today (Delespaul et al., 2016). Similarly, the patients in our study experienced outcome as an ongoing, non-linear process and not any patient regarded outcome in terms of "cure". Again, when the focus of health care and outcome shifts towards feeling empowered to tackle future challenges, this does not automatically imply that a patient ought to be symptom-free to be able to do so (Delespaul et al., 2016). In the Netherlands, some of these initiatives are being introduces in mental health care, especially for chronically and severely disturbed patients, with a particular three-fold focus: recovery, empowerment and expertise in experience. "Recovery" in this sense is re-interpreted as being able to live a meaningful life in society, despite enduring mental difficulties. Empowerment in terms of regaining self-worth is a key mechanism in doing so, as well as the strength gained from experiences which lead to self-understanding and the ability to help others (see Delespaul et al., 2016 for a description). In Belgium, similar projects

[35] https://www.who.int/about/who-we-are/constitution

arise, also particularly for those who do not fit the classical recovery-frame and health care circuit, for instance patients with substance dependency and psychotic symptomatology (Brysinck, Broekaert & Vandevelde, 2005; De Ruysscherae, Tomlinson, Vanheule & Vandevelde, 2019). Our findings deem these developments positively as they provide the field with a more hopeful perspective on mental health and more realistic role of mental health care (cf. former inpatients found resignation in more realistic expectations and ideas about treatment and recovery; study 1). Similarly, also for the treatment of major depression, authors call for a different approach to the study and conceptualization of change. They argue for a move away from a pessimistic focus on poor prognosis and chronicity, like it has been dominating the literature and consequently also the research field (see Rottenberg et al., 2018). Accordingly, our findings based on the patients who were studied in the RCT question if these changing definitions should be reserved exclusively for the more "severely disturbed" populations. Generally, it seems, shifting definitions (predominantly coming from the mental health care field) increasingly demand that also outcome research adjusts to these developments.

Implications and challenges for outcome research, clinical practice and policy

The changing understanding of mental health and health care entail important actual and future challenges for research, clinical practice and policy. In this section we will reflect on some of those we believe are pressing and promising in context of the present research. In doing so, we depart from the findings of this dissertation and thus from (and as acvocates of) a first-person approach to outcome as a starting point to elaborate on broader implications, suggestions and challenges for the field.

Outcome research departing from a broader interpretation of "clinical significance"

Our findings emphasize the importance of contextualizing quantitative outcome findings within patients' narrative and idiosyncratic context in order to interpret what pre-to-post changes signal in patients' experiences. Consequently, our findings challenge the leading

status of pre-defined outcome measurement in the field when the aim is to grasp individual-patient and clinically meaningful change. Departing from patients' perspectives, our findings support the developments of first-person definitions of clinical significance (cf. personal recovery) and focus on positive mental health. Hence, while the search for a satisfying definition of clinical significance will probably also continue among statisticians (see Wise, 2004 for instance; we recall that our discussion is not aimed at this level), we put forth a broader understanding of "clinical meaningfulness". Particularly, we can raise the question if the continued search for a definition of clinical significance (Atkins et al., 2005) should exclusively depart from a statistical interpretation and *measurement* of significance. "Significant" interpreted in terms of "having meaning, importance or salience", consequently, opens a different range of possibilities to evaluate the meaning of psychotherapy and change for patients. In that sense, the attempt to open the "black box" of psychotherapy outcome in the present dissertation by means of a three-layered analysis of the phenomenon of outcome and its clinical relevance can be considered an initial step into this process; now we will reflect on where this takes us from here

Diversification and complexification: the role of theory. A broader interpretation of "clinical significance" first of all implies that the search for "clinical meaningfulness" is answered by a *diversification* of research questions and *complexification* in the means to study them, that is: adjusted to the multi-facetted nature of change and change processes (Kazdin, 2009). The current "black box" in outcome research can be considered the result of the conventions from the past (e.g., to strictly define outcome separately from process; Kiesler, 1966). While they may have served the research aims before (e.g., the need to proof the efficacy of psychotherapy justified the reliance on standard controlled outcome studies in the 70s-90s; see Wampold & Imel, 2015), they may prevent further developments today (see also Fisher et al., 2018 on the pressing need for patient-centred research). Generally, the current state of

knowledge in the field of psychotherapy and psychology calls researchers to alter and answer more diverse and fine-grained questions (Braakmann, 2015).

Important in this endeavour, is *theory-building* research into outcome and the processes leading to therapeutic change (Kazdin, 2001b; e.g., theory-building case studies, Stiles, 2015). The applied grounded theory, data-driven analyses in the present research prove valuable qualitative means in this endeavour (cf. Elliott, 2010). Also, research into moderators (i.e., for whom and under what conditions does change occur) and mediators of change (i.e., why change occurred) that allow for the differentiation of the causal-effect interpretation of standard outcome studies, are an important added value in outcome research (Kazdin, 2009). An example can be found in the work by Johansson and colleagues (2010), who studied the mediating role of insight for the long-term effects of transference interpretation in psychodynamic treatment by making use of a dismantling design. Importantly, the researchers evaluate psychodynamic therapy outcome by means of a therapy-*specific* measures (i.e., The Psychodynamic Functioning Scales; Høgland et al., 2000) and global well-being, rather than symptom reduction. Although we can question that these designs allow grasping the multi-dimensionality of patients' holistic experiences, they address the criticized "homogeneity myths" in psychotherapy research, raised by Kiesler in 1966 but still topical today (Kazdin, 2009): as not all patients, therapists and therapies are alike, researchers should specify their questions and adjust their means of investigation to the object of study, rather than "using thermometers to weigh oranges" (see Levitt et al., 2005, p. 126 for a critical reflection). Moreover, a valuable theoretical framework to rethink simplistic linear process-outcome influences can be found in "dynamic systems theory" (see Hayes & Strauss, 1998). Based on patients' holistic experiences of therapy, we concluded that a certain factor (e.g., agency) will not always lead to successful outcome in the exact same way. Correspondingly, dynamic systems assume that change evolves over time given a certain dynamic balance between opposing processes of stability and

variability. A thorough discussion is beyond the scope of this dissertation, yet it seems to offer a valuable theoretical framework to understand why one patient achieves important changes (e.g. life changes, empowerment) over the course of therapy, while another feels "stuck between knowing and doing" and *unable* to change, and again someone else becomes overwhelmed and worsens (see Smith & Thelen, 1993).

Finding consensus in a context- and perspective-dependent definition, made explicit. For the continued search of a satisfactory indication of "clinical significance" it thus seems that, a *general* standard in terms of "convention (…) applicable to a wide variety of clinical problems (…) objective, relatively free of bias, and sound from both a psychometric and a clinical perspective" (cf. Jacobson et al., 1984), is not in line with the state of affairs in the field today. Research findings increasingly emphasize the need to address the clinical heterogeneity, call for trans-diagnostical approaches and more idiosyncratic research, as more specified knowledge and approaches are needed to help, for instance, those patients who seemingly do not benefit from treatment (see the compelling consensus statement by Bracken et al., 2012 for the field of psychiatry). Hence, calling for consensus with regard to what can be considered "clinically significant" change or outcome, we return to the observation by Strupp and Hadley (1977): "the most appropriate response must be: "It depends." It depends on what we measure, what facets of the patient's functioning we emphasize, and how we-as members of society, as patients, and as mental health professionals-choose to define a [therapy] effect" (p. 92). Generally, researchers should be more aware of the questions: *what* outcome are we measuring and *why* (i.e., for what reason/purpose)? And consequently, what is the best way to do so? For instance, when aiming to evaluate the effects of a certain therapy, one should ask whether it is *meaningful* to use symptom reduction as the outcome of interest when the treatment itself is rather concerned with interpersonal relationships and self-understanding (e.g., psychodynamic and humanistic approaches). Therapy-congruent measures in that sense can offer a valuable

alternative but are not in line with the common practice to evaluate different therapies by means of the same "thermometers" (see Levitt et al., 2005).

Moreover, researchers must consider *whom* they evaluate outcome for, which implies the acknowledgement of the inherent *perspective-dependent* value judgement and underlying standard or norm of what is considered desirable outcome (Evans & Fletcher, 2013). From a societal perspective, evaluating patients' ability to go to work may be valuable, while a patient may be aiming to get out of the house again. This broader definition hence implies that the understanding of "clinical significance" depends on the perspective through which outcome is evaluated. Thus, rather than choosing one uniform perspective that should fit all different contexts and populations, it seems more clinically useful to define the clinical significance *context- and perspective-dependent.* In accordance to broader recommendations regarding the validity in psychotherapy research (see Truijens et al., 2019), we argue that this implies that in every outcome study, researchers would have to make explicit *how* outcome was defined, *why* this definition was most suited and to *whom* or *what* it applies, hence providing the right context to interpret the findings with regard to their meaningfulness for a certain setting. The present dissertation in that sense illustrates that when the aim is to offer a patient-centred indication of the clinical meaningfulness, the definition of "personal recovery" seems more suitable than the JT classifications. Further research departing from different perspectives (e.g., therapist, society, significant others) are encouraged to evaluate similar questions, which implies an *explicit* (cf. transparent) critical reflection on whether and why the applied means can or cannot be considered useful to serve the aimed objectives (cf. Hesse-Biber, 2010, on the central influence of the research *problem* and methodology/theory in research).

Including patients' pre-treatment experiences. Finally, while a lot of emphasis in outcome research (both quantitative and qualitative) has been placed on understanding the *post*-therapy experiences of patients, little is known about patients' *pre*-level of functioning and

individual differences. Nonetheless, patients' pre-treatment characteristics are considered one of the most important predictors of outcome (Bohart & Wade, 2013). Yet these characteristics that are linked to treatment outcome, are often approached as static which thus neglects the potential of patients' position to change over time (as observed in study 1 of this dissertation). Our findings first of all highlight the need to include a thorough examination of patients *starting position* to understand the therapy process and outcome (cf. inpatients therapy trajectory differed enormously based on their previous life and health care experiences; similarly, patients in the RCT seemingly differed in their pre-treatment level of agency). As we addressed in a previous section, in the particular case of studying the effects of therapy for patients suffering from major depression, an examination of patients' depression *experiences* would be highly valuable[36]. A patient who enters treatment with the conviction of "being a depressed *person*" following a history of recurrent episodes, does not enter therapy in the same way as someone who feels like she is "losing her mind" because of "an invasion by a tiny insect that controls her thoughts". Likewise, someone who enters therapy with the explicit message (or fear) that she "*cannot* be depressed" because that would imply "looking like her father" will not engage in the same trajectory as someone who begs the therapist for a relieving diagnosis (based on excerpts from patients' narratives in the RCT, discussed in De Smet et al., unpublished manuscript). Clearly, understanding *what* psychotherapy should eventually lead to (i.e., the definition of "outcome") requires more effort on outlining *where* psychotherapy *begins* (e.g., Philips et al., 2017), in order for the trajectory in between to make sense as well. For this broader exploration, an important contribution could come from in-depth qualitative inquiry, which so far has been predominantly deployed to examine patients post-therapy experiences (McLeod, 2013).

[36] Parnas, Sass & Zahavi (2013): "The person does not manifest a series of independent referring symptoms/ signs but rather certain unities of interpenetrating experiences, feelings, expressions, beliefs, and actions, all permeated by biographical detail. These wholes and these aspects are not, in their essence, defined by a reference to underlying substrate but by their meaning." (p. 275).

What counts as "evidence" and the implications for health care decisions

A major hurdle to our suggested changes on the level of *how psychotherapy outcome is conceived*, is the remaining strong impact of the "hierarchy of evidence" on funding, policy decisions and knowledge distribution via the strict criteria listed by journals, for instance (Jadad & Enkin, 2007). These external criteria force researchers to choose a design that can meet these stringent criteria, rather than that they logically follow the need to address the most (clinically) relevant research questions. Especially, single case and qualitative in-depth studies are often restricted by the scope of the journal and page limitations. Consequently, the research that is most appealing to clinicians is the hardest to get published and thus the least evident to disseminate into the clinical field (Hyde, 2004; McLeod, 2016). The tendency to place RCTs at the top of the hierarchy is in that sense increasingly lamented, or as stated by Jadad and Enkin (2007): "fundamentally wrong. Indeed, we consider the very concept of a hierarchy of evidence to be misguided and superficial. There is no 'best evidence', except in reference to particular types of problem, in particular contexts." (p. 106).

A topical example of the consequences of this "restricted model of science" (i.e., relying on one particular source, e.g., RCT, while omitting other forms of evidence like qualitative outcome findings) for the organization of health care today can be observed in the debate surrounding the latest revision of the NICE guidelines for adult depression (i.e., the evidence-based recommendations for health and care in England). Psychotherapy researchers and clinicians across the field criticize NICE for relying entirely on clinical trials (and its strict gold standard criteria), which has important consequences for the provided therapeutic care. Firstly, the efficacy of psychotherapy is compared to medical treatments based on the criterion of "double-blind" designs (see Consort, 2010). This implies that both the person who delivers the treatment (e.g., the one who gives the pill) and the patient who receives the treatment are unaware of what that treatment is (e.g., actual medicine or placebo). As this requirement is per

definition impossible to ensure in psychotherapy (a therapist cannot be "blinded" to his own treatment and then be assumed to offer a valuable therapy), this thus places psychotherapy at a lower level on beforehand. Secondly, the guidelines are criticized for relying solely on evidence based on short-term evaluations which precludes long-term treatments form the list of effective therapies. Against what is generally recommended by researchers, that is to evaluate therapies on their *long-term* effects, outcome evaluations of one and two-year follow-up are not being considered.[37] Notwithstanding this heavy criticism on the consulted evidence, coming from the academic *and* clinical field, NICE dictates British health care and serves as the model for the development of guidelines worldwide. This topical example, however, illustrates that the seemingly appealing clear guidelines do not per se guarantee better mental health care when they are not based on the most clinically relevant research (see also Truijens, Zülke-van-Hulzen, & Vanheule, 2017).

Research developments becoming clinical applications

A particular challenge for *clinical practice* and clinicians is the incorporation of research findings and tools into daily practice, which is known as "patient-focused research (which) provides tools in order to support, but *not* replace, clinical decision-making with actual ongoing research data and specially developed decision support tools" (Castonguay, et al., 2013, p.87). In recent years, this increasingly turns towards the use of "routine outcome monitoring" (ROM). The underlying rationale is that tracking patients' progress will allow clinicians to attune treatment interventions to patients' needs. The ultimate aim is to increase the effectiveness, which is guaranteed by using continuous measurement. (Lambert, 2017). The JT method in particular, is an often-used "rationally derived method" to support therapist decisions: psychometric information based on standardized measures (e.g., the BDI) is used to make an a priori definition about a patient's status and change. Subsequently, this functions as a

[37] The stakeholder position statement on the NICE guideline for depression in adults can be retrieved via: https://cdn.ymaws.com/www.psychotherapyresearch.org/resource/resmgr/docs/downloads/StakeholderPositionStatement.pdf

benchmark for the expected change of a patient and the evaluation of the patient's progress. "The goal is, for example, to identify negative and positive developments early on in treatment and then to feed this back to therapists so they can combine science and practice immediately during the ongoing treatment" (Castonguay et al., 2013 p. 87). While this sounds promising, our findings cast doubt on the strong reliance on pre-defined measurement and statistical definitions to implement in clinical decision making, based on our findings that, in patients' experiences, this may not be that meaningful to them at all. As, while routine outcome measures can inform therapists on patients' lack of improvement and deterioration (on a certain measure), which therapists may have missed otherwise, it cannot tell the therapist how a patient interpreted the questions or what a patient meant when answering them (see chapter VI in this dissertation). Also, patients' interpretation of the items may alter throughout the course of therapy, making the comparison of the scores on the items problematic (see McLeod, 2001; Truijens, 2017). In this regard, we recall the observations by Blanton and Jaccard (2006) who addressed the issue of arbitrary metrics and repeat after Evans and Fletcher (2013) that "in the end the score on the instrument is a *measure of* something, not the something itself." (p. 32). We thus urge the need of using routine outcome monitoring as a *clinical* tool, requiring first and foremost *clinical interpretation*, rather than as a guiding tool to regulate clinical care (cf. Castonguay et al., 2013).

Secondly, our findings draw attention to on another clinical application that is informed by research practices (e.g., the use of treatment manuals; Truijens et al., 2017), namely the increasing focus on *time-limited care* which implies a predefined number of sessions, often based on patients' diagnoses and linked to reimbursement by health care insurances (Barlow, 1996). Across all studies in this dissertation, patients mentioned ongoing struggles yet remarkably, these remaining complaints did not necessarily require further therapy. As such, our patients showed that the need for therapy is not universally definable and, for instance,

depends on patients' level of agency (feeling able to cope on their own, but also, feeling up to the difficult task of therapy) and presence of a supporting social context. Moreover, this points at the inherently difficult task to predefine the number of sessions (in the literature this has been described as "the good enough level", which has been estimated at 26 sessions on average; Barkham et al., 2006). While time limits and unified treatment protocols are being implemented in clinical care, the scientific ground of time-limited therapy is still uncertain due to a scarcity in research, which currently points at both possible advantages as well as disadvantages (see De Geest & Meganck, 2019). For some of our patients, the predefined number of 20 sessions in the RCT was too limited, for others too much and for others sufficient. The patients in our study hence suggest that important patient-characteristics and idiosyncratic needs should be taken into account (apart from or rather than diagnostical information), when deciding on the duration of therapy. In other words: a first-person definition of *outcome* hence requires a compatible first-person definition of *care* (cf. Strupp & Hadley, 1977). Clearly, more research is needed to build a (clinically useful) evidence-base for these decisions (De Geest & Meganck, 2019).

Policy decisions, the tension between economical and clinical needs

A decisive role in the developments within clinical care lies in *policy decisions* that are determinant for the organization of clinical care. The current (and increasingly pressing) economic reality confronts the field with a scarcity in financial means and limited resources. As a consequence, this increases the necessity of clear and reliable judgements that ensure the best – (cost) efficient – care (McLeod, 2016). Consequently, this makes statistical straightforward definitions like the JT method highly attractive. However, our study particularly informs policy on the risks of interpreting such statistical indices in isolation from actual clinically meaningful context. Particularly in randomized controlled studies, it is assumed that the measured (average) effect is caused by the (active) treatment being evaluated, and

consequently, it is assumed that measured pre-to-post changes (for instance on the BDI as was used in our studies) reflect treatment efficacy (see Truijens et al., 2019 for a critical discussion on the underlying validity assumption). However, these mean-level statistical outcome findings merely indicate that a group, *on average* has improved, more than the control condition (on average) and more than can be expected by coincidence. Hence, they do not speak for the individual variation within the group, nor inform on what it implies for the individual patients being studied. We recall that these between-group statistics have been the impetus to define indices of clinical significance in the past (Jacobson et al., 1984; Lambert et al., 2008). However, the JT method of clinical significance is intended as an *individual*-patient account of the meaningfulness of the observed changes, and thus, interpretations regarding the effects of therapy become problematic if the individual-patient himself did not experience therapy as helpful or does not attribute change to therapy (as was found in our study). In fact, this is what Jacobson et al. (1984) cautioned for when suggesting clinical significance indices, as "for any particular subject considered in isolation from the rest of the sample, our proposed criteria do not allow one to conclude that the treatment caused the improvement." (p. 349). Nonetheless, important clinical implications are attached to these statistical findings, for instance when indices of clinical significance are synthesized in meta-analyses (cf. Westen et al., 2004). A topical example of how such interpretations are reported to policy members and the broader field can be found in the book 'Thrive', in which the authors provide – seemingly compelling – evidence for the clinical significance of CBT:

> 'The general finding is that around 50% of people treated with CBT for depression or anxiety conditions recover during treatment, and many others improve significantly. For depression, CBT is as effective as drugs in the short run, and more effective in preventing the recurrence of depression down the road. For anxiety, CBT is even more impressive. Many people with conditions like social phobia, panic disorder, or obsessive-compulsive disorder have had their condition for decades, but if successfully treated they are mostly cured for life.' '(Layard & Clark, 2014, p. 22).

A major problem with this type of conclusions is that they are provided without a clear definition of what "recovery" and "significant improvement", or broader "clinical significance" actually means in this context (see also Westen et al., 2004). Our findings warn that only unnuanced conclusions can be drawn from such decontextualized statements, with the potential consequence that it does not reflect a significant part of patients' experiences and thus will not result in better care (see Fisher et al., 2018 on the problems with generalizing group-level findings to individuals; Truijens et al., 2018). Hence, policy decisions should be informed by and based on multiple sources of data and evidence, as illustrated by Slade and Longden's review on recovery rates for psychosis (2015), which appeared far more optimistic from patients' own experiences than from the findings based on the standard outcome judgement. Likewise, the current debate overshadowing the British NICE guidelines for adult depression (cf. supra), shows the risk of limiting the provided care based on a narrow interpretation of what counts as "evidence" to construct the evidence-base of clinical care (see also Truijens et al., 2019). Hence, the pivotal role of outcome researchers should be to *inform* policy makers on the nuance that is represented *within* the data and how to interpret them, in order to allow for equally nuanced decisions and a clinically meaningful organization of clinical care that *expand* those data.

Strengths, limitations and future directions

The present dissertation offers a unique insight into standard outcome research, its findings and particularly its patient-participants, typically relied on to inform policy on the efficacy of treatments. By building onto the influential and widely-used approach to clinical significance by Jacobson and Truax (1991), our findings offer a valuable framework for interpreting other outcome studies relying on the JT classification to interpret the clinical relevance of their findings. Particularly, by looking into all four differentiated JT classifications, this dissertation fills a void in the outcome literature with regard to a more specified understanding of what the

criteria of clinically significant and reliable change imply (cf. Atkins et al., 2005; McElvaney & Timulak, 2013; Kazdin, 1999; 2001a), and hence whether the use of terms like "good outcome" and "poor outcome" is justified in light of patients' experiences. Ultimately, the present dissertation illustrates the value of studying psychotherapy outcome in a multi-layered, mixed methods way. While we do not believe the exact same findings could be found in other samples, we do value this approach to research as it enables to offer an important contextualization to the interpretation of the outcome findings in a more clinically meaningful way. Hence, the unique contribution of the present large-scale mixed methods project is that it combined different strands of research and perspectives that are most-often subject of distinct projects serving different goals, and as such allowed the creation of an internal dialogue that served both a clinical/conceptual and methodological cause (hence illustrating how 1+1=3). We encourage further developments aiming at broadening the yield of RCT-based findings by means of qualitative methodology, for instance by conducting case studies or adding a qualitative-explorative strand of research onto an existing RCT (see Midgley et al., 2014).

In light of the debate surrounding mixed methods research, one could discuss the extent to which the current project and the individual conducted studies can be regarded as *mixed*. An often-heard critique on mixed methods studies is that they do not, or not sufficiently *mix* both methods and hence cannot be identified as "mixed methods" designs. Different perspectives on what can be understood under the header of "mixed methods research" exist, however. In the present dissertation we explicitly departed from the approach by Hesse-Biber (2010) that centralizes qualitative *methodology* and an interpretative approach. The understanding of mixed methods consequently broadens, from a narrow focus on the used methods to the underlying theoretical and epistemological base as the guiding reason to make use of both quantitative and qualitative methods (see Hesse-Biber, 2010 for a description of this "comprehensive approach"). Thus, the extent to which "the findings of both methods are integrated and what

the "best" mix should be is guided by the research problem at hand" (Hesse-Biber, 2010; p. 84). Our research aimed to deepen the understanding of psychotherapy outcome and underlying change processes from the perspective of those patient-participants that are represented within the statistical outcome classifications Accordingly, the integration that was deemed most appropriate to answer our research question, was to reflect upon patients' experiences (QUAL) in comparison to the dominant statistical classifications of outcome (quan). This interpretation of mixed methods research also implies that the *mixed* nature of the project expands the use of both quantitative and qualitative data to the underlying *choices* that have been made to answer the research problem. Particularly, the quantitative data were explicitly used to define our sample of patients included into the qualitative analyses. This contrasts an exclusively qualitative standpoint that would have departed from different criteria to select patients to scrutinize patients' lived-experiences (e.g., typical or rich cases, patients' experience of outcome). Hence, the present dissertation must be regarded as a mixed methods study conducted within a qualitative-interpretative framework. Nonetheless, we acknowledge that other mixed methods approaches, departing from different research problems, would be highly useful to shed light on the phenomenon of outcome and therapeutic change (we refer the reader to Hesse-Biber, 2010 and Tashakkori & Teddlie, 2003 for different interpretations and examples of mixed methods research).

Our methodological choices inevitably impacted our findings. For instance, as we departed from the existing statistical classifications of clinical significance and studied patients' experiences in a second step, this in a way helped to keep the categories "alive". For instance, by speaking of "deteriorated patients" this entailed a possible risk to also perceive them as "deteriorated" and consequently also to understand their experiences from that "deteriorated" perspective (cf. reification). These influences were tackled as much as possible by seeking actively for both similarities and differences in patients' narratives and by representing the

findings in a nuanced manner (cf. infra). Throughout the project, different approaches to report the findings were applied and the nomenclature developed by Knox et al. (2006; see study 2; described chapter III and IV) in that sense was highly valuable to grasp both similarities and differences. Yet, whereas our approach enabled us to identify the "gap" or discordance between the statistical categories of clinical significance and patients' experiences, it requires a different approach to conclude whether the classifications make sense in patients' experiences *at all*. An important complement to our analyses would be to study patients experiences first and evaluate how and if they relate to the existing categories secondly. Our findings suggest that a different constellation of samples would probably result when relying on patients' therapy satisfaction or personal outcome experiences, yet also question whether distinct outcome categories would emerge at all (cf. personal recovery; Slade & Longen, 2015).

The applied qualitative methods in this research neither allow nor aim at generalized causal claims. Rather, they offer a unique contribution to the building of knowledge and understanding (cf. Grounded theory) on outcome and change processes that can further inform hypotheses (Kazdin, 2009), for instance, on the role of patients' agency and especially the holistic interacting role of various influencing factors. The conceptual models presented in this dissertation must thus be interpreted tentatively, although they confirm previous qualitative findings (see qualitative meta-analyses by Timulak, 2007; 2010 and Levitt et al., 2016). By integrating and synthesizing the voices of 91 patients (22 from the naturalistic study; 69 from the RCT) the present dissertation contributes a *comprehensive understanding* of outcome and change processes grounded in patients' perspectives. However, a wealth of knowledge still remains to be discovered. A key finding across our studies was the diversity and heterogeneity in patients' experiences. Given our aim to give voice to *all* patients within the selected groups, we were unable to look into these idiosyncratic differences more in detail. Further research is recommended to address the relationship between outcome measures and patients' experiences

by means of case-studies and case comparison studies to offer rich and more fine-grained analyses (see Dattilio et al., 2010; Kazdin, 2011). Also, we did not particularly focus on differences between CBT and PDT and more fine-grained analyses could focus on the patient-therapist dyad, alliance ruptures and repair and patient' satisfaction of therapy. Moreover, in the present dissertation we did not include other variables than post-therapy or follow-up interviews (i.e., retrospective inquiry of experiences). As patients' experiences of outcome may change over time, further follow-up as well as pre-treatment inquiry would be an important complement to our observations.

The choice to centralize patients' perspectives stems from the growing awareness in the field that patients are central agents (rather than passive recipients) in the therapy process and thus a vital perspective to incorporate in the evaluation of outcome and change processes (Greenberg, 1999; Levitt et al.,2016; McLeod, 2013). However, from a clinical perspective, one can argue whether patients' can fully judge their own wellbeing, progress and therapy. As many of our patients indicated, resistance and convictions of impossibility impeded the therapy process, and thus, perhaps similar biases were operating during the interviews. Patients often find it very difficult to disclose negative experiences (also observed in our studies), and although interviews (i.e., indirect from the therapist) can offer a means here, potential influences resulting from socially desirable answers cannot be excluded (McLeod, 2013). Strupp and Hadley's (1977) recommendation to study outcome from a tripartite model is supported by more recent research indicating these can all have a different idea on outcome (Altimir et al., 2018). Our project as such is limited by relying only on one perspective that consequently sheds only one particular light on the matter. We encourage future studies to create a further dialogue between different perspectives to shed various lights on the phenomenon of outcome. Therefore, we want to draw an important side note that we are not assuming that outcome research from a first-person perspective will offer us a new "gold

standard". Rather, we emphasize the role of the patient as protagonist in the therapy process, the site of change and hence an indispensable perspective to take into account (Binder et al., 2010).

Important to consider when studying the first-person perspective, is the influence of the researcher, which could be phrased as "the inevitable 'third-person' filter" when analysing patients' narratives (Stiles, 2003). By means of triangulating among different interviews and researchers, using various bracketing strategies (e.g., focused interviewing strategies, reflective diary, writing memo's), to reflect upon own interpretations, these influences were tackled as much as possible (Stiles, 2003). However, ultimately the use of member-checking (i.e., returning the interviews or qualitative analyses to participants for feedback on its content; see Birt et al., 2016 for a discussion on the use of member-checking in qualitative research) would have offered an important cross-check to our interpretations. The voice of the researcher(s) is moreover inherently present in the research project, which departed from an explicit critical stance towards some of the common practices in outcome research, as well as a conceptual understanding of outcome as an inherent complex phenomenon that requires to be studied accordingly. Likewise, also an outspoken belief in the importance of idiosyncratic experiences guided the researchers towards the study of patients' voices and hence clearly guided the research objectives and corresponding methodological choices. Yet, by making these influences explicit and substantiated, their subconscious impact was tackled as much as possible (Stiles, 2003).

The present project particularly adhered to common practices in outcome research, for instance relying on a primary psychometrically sound outcome measure, following a representative outcome research situation (De Los Reyes et al., 2011; Kazdin, 2006). However, consequently it does not allow for claims to be drawn regarding the wider field of psychotherapy research and other approaches being used to evaluate psychotherapy, which

clearly have diversified in the past decades (see Gelo, Pritz & Rieken, 2015; Lambert, 2013 for an overview). The choice of the BDI as our outcome measure moreover clearly impacted our findings, beginning with the selection of patients for our samples (cf. supra). Replication with different measures would be an interesting complement to our study, for instance selecting patients based on quality of life measures (cf. Beckstead et al., 2003) or more therapy-congruent measures (cf. Høgland et al., 2007; Levitt et al., 2005). Importantly, the addressed discrepancies between standard outcome evaluation and patients' perspectives clearly expand the classical quantitative-qualitative debate (see Ponterotto, 2005). Quantitative patient-centred outcome measures exist that for instance evaluate patients' own judgement about the efficacy of treatment. An example is the Patient Global Impression of Change Scale (PGIC; Hurst & Bolton, 2004). Patients rate their change as "a great deal better," "better, and a definite improvement," "moderately better," "somewhat better," "a little better," "almost the same," or "no change (or worse)." Other developments towards patient-centred questionnaires have been undertaken, for instance by Elliott et al. (1999) who developed the Personalized Questionnaire which departs from a thorough assessment of patients' complaints and goals which are translated into patient-specific items in a second step. These subjective outcome measures are to date no mainstream practice in psychotherapy research but could help to identify whether patients feel improved themselves (cf. reference to an internal standard for outcome evaluation). However, notwithstanding their added value to standard outcome evaluation, it does not replace the in-depth qualitative inquiry of patients' experiences which allows for the idiosyncratic understanding and contextualization. For instance, what "a little better" means to one patient may not be the case for someone else (Hill et al., 2013).

The chosen JT-method remains the most-often used method in psychotherapy outcome research today and is still regarded the method of choice in comparison to other alternatives, given its straightforward use (Ogles, 2013). Therefore, it served as an important starting point

for the present research. We acknowledge that a comparison with other approaches would have been instructive to address the wider literature and applications dealing with clinical significance. This in fact has been the focus of, for instance, Atkins et al. (2005), McGlinchey et al. (2002) and Wise (2004). Yet, while these researchers departed from questions like "does it matter which – statistical – method we use?" in the present dissertation we aimed to answer, "do our statistical methods grasp what is relevant for the patients' being studied", which thus departs from a totally different objective. In that sense we believe that the comparison of the statistical approach by Jacobson and Truax with patients' perspectives serves as an example for other statistical approaches and how they can relate to patients' experiences/perspectives, as illustrated for instance by Zimmerman and colleagues (2012) regarding the definition of recovery based on the HRSD. Similarly, we believe our findings do not limit themselves to the specific case of major depression. In the present research major depression served as the "case" to address our research questions in a representative example (both for research and clinical practice). It would be an important complement to our findings to address the issues we have raised with different approaches to define statistical clinical significance and different mental health problems, populations of patients and different context (cf. infra).

In addressing the clinical meaningfulness of outcome findings, the present research project did not focus on the broader debate regarding the difficulties voiced about outcome measurement (see McClimans, 2013). Part of our observed discrepancies between patients' experiences and pre-post symptomatic changes potentially resulted from validity issues, self-report bias and measurement errors (see Desmet, 2018). We refer the reader to the work by Truijens (2016) for a critical discussion on these issues. We also did not discuss the more general debate surrounding the limitations of highly standardized outcome research (cf. RCTs; cf. Westen et al., 2004). Despite the inevitable selected group of patients willing and eligible to participate in our research (e.g., excluding patients with psychosis and primary substance

abuse), our samples did present, for instance, comorbid problems in line with clinical reality. Also, our observations show interesting resemblances to what has been observed in naturalistic samples (Ekroll & Rønnestad, 2015). Importantly, our central aim was to provide insight into those perspectives of the patient-participants *within* an outcome study, as a means to offer a clinically meaningful understanding of the measured changes in an often-criticized research setting. In doing so, our findings convincingly show that the patient-participants included in an RCT are not necessarily "ideal" or "pure" patients undone from all clinical complexity, as the often-heard critique on the eligibility criteria in RCTs states (cf. Westen et al., 2004). While we acknowledge the impact of excluding and including patients based on certain criteria, our findings more compellingly point at the variety in experiences that lays beyond the shared inclusion criterion of major depression (apart from age, language, region etc.). Thus, turning the remark around, the most important problem with regard to clinical relevance is not the *patients* who are included in the RCT, but how they are subsequently described and represented in outcome research and findings.

Moreover, including both a naturalistic and standardized study in one and the same project, allowed us to broaden the perspective and build our knowledge based on patients from distinct settings. When comparing these seemingly distinct populations (which can generally be differentiated on their level of chronicity and severity of difficulties and treatment intensity), remarkable similarities were found in the domains of experienced change. An important contribution to our search for understanding outcome and its clinical meaningfulness, however, lays in expanding the discussion on clinical significance to those contexts in which the JT method is ill-suited. For instance, particularly chronically ill patients or patients who score below the clinical cut-off prior to engaging in therapy are (practically and per definition) unable to reach "clinically significant change". While it is clear they *can* achieve significant changes, a different approach and workable definition is needed to get a hold of this (Lambert et al.,

2008). Hence, in general, but perhaps especially in these more complex clinical populations, the continued search for a useful conceptualization and evaluation of treatment outcome and its clinical meaningfulness is essential.

Conclusion

Psychotherapy outcome research has a major influence on clinical practice as it informs policy decisions on the organization of mental health care. Useful research findings that allow for a clinically meaningful interpretation and application are thus vital (McLeod, 2016). The findings of this dissertation challenge the assumption that the widely-used statistical indices of clinical significance (by Jacobson & Truax, 1991) are sufficient to offer researchers, clinicians and policy members these clear and useful conclusions. Apart from a rough differentiation between those patients who clearly improved (i.e., show nearly-complete symptom reduction) and those who clearly worsened (i.e., show a substantial increase in symptom severity) it showed unjustified to assemble patients into groups, given the striking within-sample differences and between-sample similarities. Moreover, the statistical indices of clinical significance failed to grasp patients' holistic experience of multi-dimensional change, nor its ongoing non-linear character. From patients' perspectives, it showed that the effects of psychotherapy need to be interpreted in light of patients' level of agency, contextual demands and the patient-therapy/therapist dyad across time, and thus challenging static or linear process-outcome claims. Hence, we conclude that in order to be of clinical use, outcome researchers should adjust their means to the inherent complexity of the phenomenon rather than constraining its understanding by means of statistical convention. From a first-person perspective, any definition of clinical significance should depart from an internally rather than externally defined criterion, in correspondence to the idiosyncratic phenomenological experience of mental difficulties like depression and in line with more recent conceptual developments of positive health and personal recovery. As such, the search for clinical meaningfulness in outcome

research continues and hence can be considered an ongoing (perhaps non-linear) process, like the outcome experiences expressed by patients in this study. In the continuing quest for clinical relevance in outcome research we therefore call for a shift towards a broader understanding of the notion of clinical "significance", beyond a narrow statistical interpretation. This implies the critical considerations of *what* outcome is measured *why* and for w*hom* or *what* purpose, in order to provide a contextualized and perspective-dependent interpretation of the clinical meaningfulness of findings *within* each outcome study. Integrating different perspectives and the use of multiple methods in line with the multi-facetted nature of the phenomenon of outcome are important avenues to follow. Ultimately, we should arrive at a nuanced and contextualized foundation of evidence that effectuates a clinically meaningful transition of findings to policy, clinical practice and patients.

References

Altimir, C., Krause, M., de la Parra, G., Dagnino, P., Tomicic, A., Valdés, N., … Vilches, O. (2010). Clients', therapists', and observers' agreement on the amount, temporal location, and content of psychotherapeutic change and its relation to outcome. *Psychotherapy Research, 20*(4), 472–487. https://doi.org/10.1080/10503301003705871

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Arlington, VA: American Psychiatric Association Publishing.

Atkins, D. C., Bedics, J. D., Mcglinchey, J. B., & Beauchaine, T. P. (2005). Assessing clinical significance: Does it matter which method we use? *Journal of Consulting and Clinical Psychology*, *73*(5), 982–989. https://doi.org/10.1037/0022-006X.73.5.982

Barber, J.P., Muran, J.C., McCarthy, K.S. & Keefe, J. (2013). Research on dynamic therapies. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behaviour change* (6th Ed.) (pp. 443-494). Hoboken, New Jersey: John Wiley & Sons, Inc.

Barkham, M., Connell, J., Stiles, W. B., Miles, J. N. V., Margison, F., Evans, C., & Mellor-Clark, J. (2006). Dose-effect relations and responsive regulation of treatment duration: The good enough level. *Journal of Consulting and Clinical Psychology, 74*(1), 160–167. https://doi.org/10.1037/0022-006X.74.1.160

Barlow, D. H. (1996). Health Care Policy, Psychotherapy Research, and the Future of Psychotherapy. *American Psychologist*, 9.

Beck A.T., Steer R.A., & Brown G.K. (1996). *Manual for the Beck Depression Inventory II.* San Antonio, TX: Psychological Corporation

Beckstead, D. J., Hatch, A. L., Lambert, M. J., Eggett, D. L., Goates, M. K., & Vermeersch, D. A. (2003). Clinical significance of the Outcome Questionnaire (OQ-45.2). *The Behavior Analyst Today, 4*(1), 86. https://doi.org/10.1037/h0100015

Binder, P.-E., Holgersen, H., & Nielsen, G. H. (2010). What is a "good outcome" in psychotherapy? A qualitative exploration of former patients' point of view. *Psychotherapy Research, 20*(3), 285–294. https://doi.org/10.1080/10503300903376338

Birt, L., Scott, S., Cavers, D., Campbell, C., & Walter, F. (2016). Member Checking: A Tool to Enhance Trustworthiness or Merely a Nod to Validation? *Qualitative Health Research*, *26*(13), 1802–1811. https://doi.org/10.1177/1049732316654870

Blanton, H., & Jaccard, J. (2006). Arbitrary metrics in psychology. *The American Psychologist*, *61*(1), 27–41. doi: 10.1037/0003-066X.61.1.62

Braakmann, D. (2015). Historical paths in psychotherapy research. In O. C. G. Gelo, A. Pritz & B. Rieken (Eds.), *Psychotherapy research: foundations, process and outcome* (pp. 39-66). Vienna, Austria: Springer.

Bohart, A. (1993). Experiencing: The basis of psychotherapy. Journal of Psychotherapy Integration, 3, 51-67.

Bohart, C. A., & Wade, A. M. (2013). The client in psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behaviour change* (6th Ed.) (pp. 219-257). Hoboken, New Jersey: John Wiley & Sons, Inc.

Bordin, E. (1979). The generalizability of the psychoanalytic concept of the working alliance. Psychotherapy: *Theory, Research and Practice, 16*, 252–260. doi:10.1037/h0085885

Bracken, P. et al. (2012). Psychiatry beyond the current paradigm. *British Journal of Psychiatry, 201*, 430-434.

Brysinck, D., Broekaert, E, & Vandevelde, S. (2005). Psychosis and newly emerging dependency groups: the search for an adapted model of care at the Psychiatric Centers Sleidinge in Belgium. *Therapeutic communities, 26*(2)

Burlingame, Strauss, & Joyce (2013). Change mechanisms and effectiveness of small group treatments. In M. J. Lambert (Ed.), Bergin and Garfield's handbook of psychotherapy and behaviour change (6th Ed.) (pp. 443-494). Hoboken, New Jersey: John Wiley & Sons, Inc.

Carey, T. A., Carey, M., Mullan, R. J., Murray, L. K., & Spratt, M. B. (2006). Psychological Change: What changes and how does it occur? A critical review. *Counselling Psychology Review, 21*(4) 28-38

Carey, T. A., Carey, M., Stalker, K., Mullan, R. J., Murray, L. K., & Spratt, M. B. (2007). Psychological change from the inside looking out: A qualitative investigation. *Counselling and Psychotherapy Research*, *7*(3), 178–187. https://doi.org/10.1080/14733140701514613

Castonguay, L. G., & Hill, C. E. (Eds.). (2012). *Transformation in psychotherapy: Corrective experiences across cognitive behavioral, humanistic, and psychodynamic approaches.* Washington, DC, US: American Psychological Association.

Castonguay, L., Barkham, M., Lutz, W., & McAleavey, A. (2013). Practice-oriented research: Approaches and applications. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behaviour change* (6th Ed.) (pp 85-133). Hoboken, New Jersey: John Wiley & Sons, Inc.

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 behaviour change* (5th Ed.) (pp. 194-226). Hoboken, New Jersey: John Wiley & Sons, Inc.

Consort (2010). CONSORT 2010 checklist of information to include when reporting a randomised trial. Retrieved from www.consort-statement.org.

Cuijpers, P., Huibers, M, Ebert, D. D. Koole, S.L., & Gerhard Andersson, G (2013). How much psychotherapy is needed to treat depression? A meta-regression analysis. Journal of Affective Disorders, 149, 1-13, https://doi.org/10.1016/j.jad.2013.02.030

Dattilio, F. M., Edwards, D. J. A., & Fishman, D. B. (2010). Case studies within a mixed methods paradigm: Toward a resolution of the alienation between researcher and practitioner in psychotherapy research. *Psychotherapy: Theory, Research, Practice, Training*, *47*(4), 427–441. https://doi.org/10.1037/a0021181

De Geest, R. M., & Meganck, R. (2019). How do time limits affect our psychotherapies? A literature review. *Psychologica Belgica.*

Dehue, T. (2014). *Betere mensen. Over gezondheid als keuze en koopwaar.* Nederland: Atlas Contact

Delespaul, P., Milo, M., Schalken, F., Boevink, W., & van Os, J. (2016). *Goede GGZ! Nieuwe concepten, aangepaste taal en betere organisatie.* Netherlands: Diagnosis publishers

Delgado-Romero, E. A., & Romero-Shih, A. I. (2016). Patricia Arredondo: Creating a Pathway for Cultural Empowerment. *The Counseling Psychologist*, *44*(8), 1212–1235. https://doi.org/10.1177/0011000016683943

De Los Reyes, A., Kundey, S. M. A., & Wang, M. (2011). The end of the primary outcome measure: A research agenda for constructing its replacement. *Clinical Psychology Review*, *31*(5), 829–838. https://doi.org/10.1016/j.cpr.2011.03.011

De Ruysscher, C., Tomlinson, P., Vanheule, S., & Vandevelde, S. (2019). Questioning the professionalization of recovery: A collaborative exploration of a recovery process. *Disability & Society*, 1–22. https://doi.org/10.1080/09687599.2019.1588708

De Smet, M. M., & Meganck, R. (2018). Understanding Long-term Outcome from the Patients' Perspective: A Mixed Methods Naturalistic Study on Inpatient Psychotherapy. *Psychologica Belgica, 58*(1), 276. doi: 10.5334/pb.432

De Smet, M. M., Meganck, R., Vannieuwenhove, K., Truijens, F. L., & Desmet, M. (2019). No Change? A Grounded Theory Analysis of Depressed Patients' Perspectives on Non-improvement in Psychotherapy. *Frontiers in Psychology, 10*(588). doi: 10.3389/fpsyg.2019.00588

De Smet, M.M., Cornelis, S., Acke, E., Reitske, M., Truijens, F., & Desmet, M. (2019). Making Sense of Deterioration in Psychotherapy: An Interpretative Phenomenological Analysis of Depressed Patients' Perspectives (manuscript forthcoming)

De Smet, M.M. et al. (n.d.). The patients-participant's perspective on major depression: A qualitative analysis of how RCT participants experience their complaints (unpublished manuscript).

Desmet, M. (2018). *The Pursuit of Objectivity in Psychology*. Gent: Borgerhoff & Lamberigts.

Dimidjian, S., & Hollon, S. D. (2010). How would we know if psychotherapy were harmful? *American Psychologist*, *65*(1), 21–33. http://dx.doi.org/10.1037/a0017299

Drisko, J. (2004). Common factors in psychotherapy outcome: Meta-analytic findings and their implications for practice and research. Families in Society. *The Journal of Contemporary Social Services, 85*(1), 81-90. Retrieved from http://journals.sagepub.com

Ekroll, V. B., & Rønnestad, M. H. (2017). Processes and changes experienced by clients during and after naturalistic good-outcome therapies conducted by experienced psychotherapists. *Psychotherapy Research*, *27*(4), 450–468. https://doi.org/10.1080/10503307.2015.1119326

Elliott, R. (2010) Psychotherapy change process research: Realizing the promise. *Psychotherapy Research, 20*(2), 123-135, DOI: 10.1080/10503300903470743

Elliott, R., Mack, C., & Shapiro, D. (1999). *Simplified Personal Questionnaire Procedure*. Toledo, OH: University of Toledo

Evans, I. M., & Fletcher, A. (2013). Focus on change: implications for the understanding and evaluation of psychological interventions. *Psychologia, 56*(2), 113–130. https://doi.org/10.2117/psysoc.2013.113

Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences*, *115*(27), E6106–E6115. https://doi.org/10.1073/pnas.1711978115

Foster, S. L., & Mash, E. J. (1999). *Assessing Social Validity in Clinical Treatment Research: Issues and Procedures*. 12.

Gianakis, M., & Carey, T. A. (2008). A review of the experience and explanation of psychological change. *Counselling Psychology Review, 23*(3), 27-38.

Greenberg, L. & Pinsof, W.M. (1986). Process research: current trends and future perspectives. In L. Greenberg and W. M. Pinsof (Eds) *The psychotherapeutic process. A research handbook* (pp.3-20). New York: NY, The Guilford Press

Greenberg, L. S. (1999). Ideal psychotherapy research: A study of significant change processes. *Journal of Clinical Psychology*, *55*(12), 1467–1480.

Gupta, M. (2007). Does evidence-based medicine apply to psychiatry? *Theoretical Medicine and Bioethics*. DOI 10.1007/s11017-007-9029-x

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.

Hayes, A. M., & Strauss, J. L. (1998). Dynamic Systems Theory as a Paradigm for the Study of Change in Psychotherapy: An Application to Cognitive Therapy for Depression. *Journal of consulting and clinical psychology, 66*(6), p. 939-947

Hesse-Biber, S. (2015). Mixed Methods Research: The "Thing-ness" Problem. *Qualitative Health Research*, *25*(6), 775–788. https://doi.org/10.1177/1049732315580558

Hill, C. E., Knox, S., Thompson, B. J., Williams, E. N., Hess, S. A., & Ladany, N. (2005). Consensual qualitative research: An update. *Journal of Counseling Psychology, 52,* 196–205. doi:10.1037/0022-0167.52.2.196

Hill, C. E., Chui, H., & Baumann, E. (2013). Revisiting and reenvisioning the outcome problem in psychotherapy: An argument to include individualized and qualitative measurement. *Psychotherapy, 50*(1), 68–76. doi: 10.1037/a0030571

Høglend, P., Bøgwald, K.-P., Amlo, S., Heyerdahl, O., Sørbye, Ø., Marble, A., . . . Bentsen, H. (2000). Assessment of change in dynamic psychotherapy. *Journal of Psychotherapy Practice and Research, 9,* 190–199.

Huber, M., Knottnernus, J. A., Green, L., van der Horst, V., Jadad, A. R., Kromhout, D., … Smid, H. (2011). How should we define health? *British Medical Journal,* 1–3. doi: 10.1136/bmj.d4163

Hurst, H., & Bolton, J. (2004). Assessing the clinical significance of change scores recorded on subjective outcome measures. *Journal of Manipulative and Physiological Therapeutics*, *27*(1), 26–35. https://doi.org/10.1016/j.jmpt.2003.11.003

Hyde, P. (2004). Fool's gold: Examining the use of gold standards in the production of research evidence. *British Journal of Occupational Therapy, 67*, 89-94.

Hyman, S.E. (2010). The diagnosis of mental disorders: the problem of reification. *Annual review of clinical psychology, 6*, 155-179.

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

Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy*, *15*(4), 336–352. https://doi.org/10.1016/S0005-7894(84)80002-7

Jadad, A., & Enkin, M. (2007). Randomized controlled trials: Questions, answers, and musings (2nd ed.). Malden, MA: Blackwell Publishing.

Jennissen, S., Huber, S., Ehrenthal, J. C., Schauenburg, H., & Dinger, U. (2018) Association Between Insight and Outcome of Psychotherapy: Systematic Review and Meta-Analysis. American Journal of Psychiatry. DOI: 10.1176/appi.ajp.2018.17080847

Johansson, P., Høglend, P., Ulberg, R., Amlo, S., Marble, A., Bøgwald, K.-P., … Heyerdahl, O. (2010). The mediating role of insight for long-term improvements in psychodynamic therapy. *Journal of Consulting and Clinical Psychology*, *78*(3), 438–448. https://doi.org/10.1037/a0019245

Joiner, T. E., Jr. (2000). Depression's vicious scree: Self-propagating and erosive processes in depression chronicity. *Clinical Psychology: Science and Practice*, *7*(2), 203–218. doi:10.1093/clipsy/7.2.203

Kazdin, A. E. (1977). Assessing the clinical or applied importance of behavior change through social validation. *Behavior modification*, *1*(4), 427-452.

Kazdin, A. E. (1999). The meanings and measurement of clinical significance. *Journal of Consulting and Clinical Psychology*, 332–339

Kazdin, A. E. (2001a). Almost clinically significant (p<.10): Current measures may only approach clinical significance. *Clinical psychology: Science and practice,* 8(4), 455-462. https://onlinelibrary.wiley.com /doi/pdf/10.1093/clipsy.8.4.455

Kazdin, A. E. (2001b). Bridging the Enormous Gaps of Theory With Therapy Research and Practice. *Journal of Clinical Child & Adolescent Psychology*, *30*(1), 59–66. https://doi.org/10.1207/S15374424JCCP3001_7

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

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

Kazdin, A. E. (2011). *Single-case research designs: methods for clinical and applied settings (2nd Ed.)* Oxford, England: Oxford University Press

Kiesler, D. J. (1966). Some myths of psychotherapy research and the search for a paradigm. *Psychological Bulletin, 65*(2), 110–136.

Kiesler, D. J. (1971). Patient experiencing level and successful outcome in individual psychotherapy of schizophrenics and psychoneurotics. Journal of consulting and clinical psychology, 37, 370-385

Kiesler, D. J. (1973). The process of psychotherapy. Chicago: Aldine

Knox, S., Schlosser, L. Z., Pruitt, N. T., & Hill, C. E., (2006). A qualitative examination of graduate advising relationships: The advisor perspective. *The Counseling Psychologist, 34,* 489–518. doi:10.1177/0011000006290249

Lambert, M. J. (1992). Psychotherapy outcome research: Implications for integrative and eclectic therapists. In J. C. Norcross & M. R. Goldfried (Eds.), *Handbook of psychotherapy integration* (pp. 94–129) New York: Basic Books.

Lambert, M. J., Hansen, N. B., Umphress, V., Lunnen, K., Okiishi, J., Burlingame, G. M., & Reisinger, C. W. (1996). *Administration and scoring manual for the OQ-45.2*. East Setauket, NY: American Professional Credentialing Services.

Lambert, M. J., Hansen, N. B., & Bauer, S. (2008). Assessing the Clinical Significance of Outcome Results. In A. M. Nezu & C. M. Nezu (Eds.). *Evidence-based outcome research. A Practical Guide to Conducting Randomized Controlled Trials for Psychosocial Interventions* (pp. 359-378). New York, NY: Oxford University Press

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. DOI:10.1080/10503300902849483

Lambert, M. J. (2013). The efficacy and effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behaviour change* (6th Ed.) (pp. 169-218). Hoboken, New Jersey: John Wiley & Sons, Inc.

Lambert, M. J. (2017). Maximizing Psychotherapy Outcome beyond Evidence-Based Medicine. *Psychotherapy and Psychosomatics*, *86*(2), 80–89. https://doi.org/10.1159/000455170

Lampropoulos, G. K. (2011). Failure in psychotherapy: An introduction. Journal of Clinical Psychology: In Session, 67(11), 1093–1095. doi:10.1002/jclp.20858

Layard, R., & Clark, D. M. (2014). *Thrive. How Better Mental Health Care Transforms Lives and Saves Money.* Allen Lane, Penguin Group, London

Levitt, H. M., Stanley, C. M., Frankel, Z., & Raina, K. (2005). An Evaluation of Outcome Measures Used in Humanistic Psychotherapy Research: Using Thermometers to Weigh Oranges. *The Humanistic Psychologist*, *33*(2), 113–130. https://doi.org/10.1207/s15473333thp3302_3

Levitt, H. M., Pomerville, A., & Surace, F. I. (2016). A qualitative meta-analysis examining clients' experiences of psychotherapy: A new agenda. *Psychological Bulletin*, *142*(8), 801–830.

Malterud, K. (2001). The art and science of clinical knowledge: Evidence beyond measures and numbers. *The Lancet*, 358, 397-340.

McAdams, D. P. (2001). *Coding Autobiographical Episodes for Themes of Agency and Communion*.

McClimans, L. (2013). The role of measurement in establishing evidence. *Journal of Medicine and Philosophy*, *38*, 520–538.

McElvaney, J., & Timulak, L. (2013). Clients' experience of therapy and its outcomes in 'good' and 'poor' outcome psychological therapy in a primary care setting: An exploratory study. *Counselling and Psychotherapy Research*, *13*(4), 246–253. https://doi.org/10.1080/14733145.2012.761258

McGlinchey, J. B., Atkins, D. C., & Jacobson, N. S. (2002). Clinical significance methods: Which one to use and how useful are they? *Behavior Therapy*, *33*(4), 529–550. https://doi.org/10.1016/S0005-7894(02)80015-6

McLeod, J. (2013). Qualitative research. Methods and contributions. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed.) (pp. 49-84). New York, NY: John Wiley & Sons.

McLeod, J. (2016). *Using research in counselling and psychotherapy.* London, England: sage

Meganck, R., et al. (2017). The Ghent Psychotherapy Study (GPS) on the differential efficacy of supportive-expressive and cognitive behavioural interventions in dependent and self-critical depressive patients: study protocol for a randomized controlled trial. *Trials*, *18*(126), 1-11. DOI:10.1186/s13063-017-1867-x

Midgley, N., Ansaldo, F., & Target, M. (2014). The meaningful assessment of therapy outcomes: Incorporating a qualitative study into a randomized controlled trial evaluating the treatment of adolescent depression. *Psychotherapy*, *51*(1), 128-137. Doi:10.1037/a0034179

Miller, S. D., Duncan, B. L., & Hubble, M. A. (1997). Escape from Babel: Toward a unifying language for psychotherapy practice. New York: Norton.

Mohr, D. C., (1995). Negative outcome in psychotherapy: A critical review. *Clinical Psychology: Science and Practice, 2*(1), 1– 27. doi:10.1111/j.1468-2850.1995.tb00022.x

Nilsson, T., Svensson, M., Sandell, R., & Clinton, D. (2007). Patients' experiences of change in cognitive-behavioural therapy and psychodynamic therapy: a qualitative comparative study. *Psychotherapy Research*, *20*, 37-41. doi: 10.1080/10503300601139988

Norcross, J. C., Krebs, P. M., & Prochaska, J. O. (2011). Stages of change. In J. C. Norcross (Ed.), *Psychotherapy relationships that work: Evidence-based responsiveness* (2nd ed., pp. 279–300). New York, NY: Oxford University Press.

Ogles (2013). Measuring change in psychotherapy research. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behaviour change* (6th Ed.) (pp. 134-166). Hoboken, New Jersey: John Wiley & Sons, Inc.

Parnas, J., Sass, L. A., & Zahavi, D. (2013). Rediscovering psychopathology: the epistemology and phenomenology of the psychiatric object. *Schizophrenia Bulletin, 39*(2), 270-277.

Philips, B., Karlsson, R., Nygren, R., Rother-Schirren, A., & Werbart, A. (2018). Early therapeutic process related to dropout in mentalization-based treatment with dual diagnosis patients. *Psychoanalytic Psychology*, *35*(2), 205–216. https://doi.org/10.1037/pap0000170

Ponterotto, J. G. (2005). Qualitative research in counseling psychology: A primer on research paradigms and philosophy of science. *Journal of Counseling Psychology*, *52*(2), 126–136. https://doi.org/10.1037/0022-0167.52.2.126

Ratcliffe, M. (2014). *Experiences of depression: A study in phenomenology*. OUP Oxford.

Ross, E. C., Polaschek, D. L. L., & Ward, T. (2008). The therapeutic alliance: A theoretical revision for offender rehabilitation. *Aggression and Violent Behavior, 13*(6), 462–480. doi:10.1016/j.avb.2008.07.003

Rottenberg, J., Devendorf, A. R., Kashdan, T. B., & Disabato, D. J. (2018). The Curious Neglect of High Functioning After Psychopathology: The Case of Depression. *Perspectives on Psychological Science*, *13*(5), 549–566. https://doi.org/10.1177/1745691618769868

Safran, J. D., Muran, J. C., & Eubanks-Carter, C. (2011). Repairing alliance ruptures. *Psychotherapy*, *48*(1), 80–87. https://doi.org/10.1037/a0022140

Slade (2009). *Personal Recovery and Mental Illness: A Guide for Mental Health Professionals.* Cambridge: Cambridge University Press.

Slade, M., & Longden, E. (2015). Empirical evidence about recovery and mental health. *BMC Psychiatry*, *15*(1). https://doi.org/10.1186/s12888-015-0678-4

Smith, L. B., & Thelen, E. (Eds.). (1993). *A Dynamic systems approach to development: Applications*. Cambridge, Mass: MIT Press.

Stiles, W. B. (1993). Quality Control in Qualitative Research. *Clinical psychology Review, 13,* 593-618. doi: 10.1016/0272-7358(93)90048-Q

Stiles, W. B. (2015). Theory-building, enriching, and fact-gathering: Alternative purposes of psychotherapy research. In O. Gelo, A. Pritz, & B. Rieken (Eds.), *Psychotherapy research: General issues, process, and outcome* (pp. 159-179). New York: Springer-Verlag.

Strupp, H. H., & Hadley, S. W. (1977). A tripartite model of mental health and therapeutic outcome: With specific reference to negative effects in psychotherapy. *American Psychologist, 32,* 187–196. doi:10.1037/0003- 066X.32.3.187

Tashakkori, A., & Teddlie, C. (Eds.). (2003). *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage.

Timulak, L. (2007). Identifying core categories of client-identified impact of helpful events in psychotherapy: A qualitative meta-analysis. *Psychotherapy Research*, *17*, 310–320. https://doi.org/10.1080/10503300600608116

Timulak, L. (2010). Significant events in psychotherapy: An update of research findings. *Psychology and Psychotherapy: Theory, Research and Practice*, *83*(4), 421–447. https://doi.org/10.1348/147608310X499404

Tingey, R. C., Lambert, M. J., Burlingame, G. M., & Hansen, N. B. (1996). Assessing clinical significance: Proposed extensions to method. *Psychotherapy Research, 6*, 109–123.

Truijens, F. L. (2017). Do the numbers speak for themselves? A critical analysis of procedural objectivity in psychotherapeutic efficacy research. *Synthese*, *194*(12), 4721–4740. https://doi.org/10.1007/s11229-016-1188-8

Truijens F.L., Cornelis, S, Desmet, M., De Smet M. M., & Meganck, R. (2019) Validity Beyond Measurement: Why Psychometric Validity Is Insufficient for Valid Psychotherapy Research. *Frontiers in Psychology,* 10:532. doi: 10.3389/fpsyg.2019.00532

Truijens, F., Zülke-van-Hulzen, L. Z., & Vanheule, S. (2019). To manualize, or not to manualize: Is that still the question? A systematic review of empirical evidence for manual superiority in psychological treatment. *Journal of Clinical Psychology*, *75*(3), 329–343. https://doi.org/10.1002/jclp.22712

Tryon, W. W. (2005). Possible mechanisms for why desensitization and exposure therapy work. *Clinical Psychology Review, 25,* 67- 95.

van der Does, A.J.W. (2002). *BDI-II-NL. Handleiding. De Nederlandse versie van de Beck Depression Inventory-2nd edition.* Lisse: Harcourt Test Publishers.

van Os, J., Guloksuz, S., Vijn, T. W., Hafkenscheid, A., & Delespaul, P. (2019). The evidence-based group-level symptom-reduction model as the organizing principle for mental health care: Time for change?: The evidence-based group-level symptom-reduction model as the organizing principle for mental health care: time for change? *World Psychiatry*, *18*(1), 88–96. https://doi.org/10.1002/wps.20609

Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work*. New York: Routledge.

Watzlawick, P., Beavin, J., & Jackson, D. (1967). *Pragmatics of human communication.* New York: Norton

Westen, D., Novotny, C. M., & Thompson-Brenner, H. (2004). The Empirical Status of Empirically Supported Psychotherapies: Assumptions, Findings, and Reporting in Controlled Clinical Trials. *Psychological Bulletin, 130*(4), 631–663. doi: 10.1037/0033-2909.130.4.631

Wise, E. A. (2004). Methods for Analyzing Psychotherapy Outcomes: A Review of Clinical Significance, Reliable Change, and Recommendations for Future Directions. *Journal of Personality Assessment*, *82*(1), 50–59. https://doi.org/10.1207/s15327752jpa8201_10

World Health Organization (2017). Depression and other common mental disorders: Global health estimates. Retrieved from http://apps.who.int/iris/bitstream/10665/254610/1/WHO-MSD-MER-2017.2

Zimmerman, M., Martinez, J. A., Attiullah, N., Friedman, M., Toba, C., Boerescu, D. A., & Rahgeb, M. (2012). Why do some depressed outpatients who are in remission according to the Hamilton Depression Rating Scale not consider themselves to be in remission? *Journal of Clinical Psychiatry*, *73*(6), 790-795. https://doi.org/10.1002/da.21987

English summary

Opening the black box of psychotherapy outcome: A mixed methods study on patients' perspectives

In this dissertation, we address the complicated relationship between academic research and clinical practice by focusing on the *clinical usefulness* of findings from outcome research. In doing so, we particularly (yet not exclusively) address the implications for RCT-based outcome findings, given their status as gold standard research and consequential important influence on the organization of the health care field. We departs from the influential statistical definition of "clinical significance", as developed by Jacobson and Truax (the "JT method", 1991). To date this method is considered the "method of choice" to indicate the clinical relevance of statistical outcome findings. Based on a twofold evaluation (i.e., reliable change and evolution from functional to dysfunctional range) patients' can be marked as "recovered", "improved", "unchanged" or "deteriorated". This method provides an individual-patient estimate of the relevance of the measured pre-to-post change and as such provides insight into the variation *within* general outcome findings. Notwithstanding the important contribution to the

interpretation of statistical mean-level results, in this dissertation we depart from two crucial questions that remain unresolved in the "black box" of psychotherapy outcome. Firstly, what does "clinically significant change" mean in light of patients' own experiences? Secondly, which change processes underly clinically significant outcome results? The pre-post, one-dimensional outcome measurement included in the JT procedure does not allow answering these questions, hence calling for further conceptual and methodological developments. In the present dissertation we therefore examine the perspectives of those patients-participants included in psychotherapy research and thus those voices relied on to conclude on the outcome of psychotherapy. To address our objectives, we present a representative case of the type of research findings that inform clinical practice.

Three main objectives guided the inquiries: 1) scrutinizing patients' experiences of change, which allows to deepen the understanding of "psychotherapy outcome" from a first-person perspective, 2) elucidating the change processes patients believe have influenced meaningful change, which enables to explore underlying change mechanisms, and 3) examining how well the pre-defined classification by Jacobson and Truax (1991) allows capturing meaningful change in light of patients' own perspectives. Hence, both a clinical/conceptual and methodological question underly this research project. In order to answer our objectives, we put forward a three-layered, mixed methods approach, following a sequential explanatory design as described by Hesse-Biber (2010). Specifically, we depart from the statistical (quantitative) outcome evaluation in terms of clinical significance (i.e., descriptive layer), and expand this pre-post outcome evaluation by means of in-depth qualitative inquiry of patients' experienced changes (i.e., explorative layer) and the change processes they believe have influenced important changes (i.e., explanatory layer). Ultimately, this allows to open the existing general outcome categories and shed light on the perspectives of those patients represented within.

Four mixed methods studies were conducted in order to answer the central research objectives; they are presented in five consecutive chapters. Study 1 (chapter II) presents a naturalistic study in an inpatient setting in Flanders (Belgium). Long-term outcome was studied by enriching quantitative outcome evaluation with patients' perspectives five to six years after ending treatment. Using a data-driven thematic analysis, the client change interview of 22 former patients were examined. The resulting tentative conceptual model of long-term outcome is presented and discussed. The study serves as the pilot study for the rest of the project, which is situated within an RCT on the treatment of major depression (the Ghent Psychotherapy Study; Meganck et al., 2017). In study 2 (chapter III and IV) we examine the positive end of the outcome continuum. Patients showing "good" outcome, i.e., recovery and improvement in depression symptom severity, as measured by means of the Beck Depression Inventory (BDI), were included. A grounded theory analysis on the interviews with 28 recovered and 19 improved patients was conducted to scrutinize patients' experienced change over the course of psychotherapy (chapter III) and the change processes or influencing factors they believe have led to these changes (chapter IV). Subsequently, in study 3 and 4 we address these questions at the negative end of the outcome continuum. A grounded theory analysis of the narratives of the 19 patients who showed no reliable change in symptom scores, resulted in a conceptual model of "non-improvement" (study 3; chapter V). An interpretative phenomenological analysis (IPA) was finally conducted on three cases marked as "deteriorated" based on their outcome scores (study 4; chapter VI). In all chapters we reflect on the relationship between the statistical outcome evaluation on the one hand and patients' experiences on the other, which ultimately allows to draw conclusions regarding the usefulness of standard outcome research for the transference of findings to clinical care.

In the general discussion (chapter VII) a thorough integration of the separate findings of all studies is presented, resulting in an integrative conceptual model of outcome. In answer

to our first main objectives we found that, from patients' perspectives outcome is an ongoing, non-linear process characterized by individual differences. Patients moreover described a holistic experience of multi-dimensional change. Differences among patients' experiences were observed with regard to the level of empowerment, internalization of insights and reaching intra-and/or interpersonal balance. Moreover, also the extent to which therapeutic tools and insight could be translated into actual life or behavioural changes and whether ongoing struggles were perceived as normal or impeding and required further treatment or not differed among patients.

The answer to our second objective allowed to interpret these experienced changes and observed individual differences, leading to tentative explanations. Particularly, our findings point at multiple change processes – inside and outside therapy – interacting over time, which reflects a holistic psychotherapy experience, imbedded in patients' life, person and context. This corresponds to patients' holistic experience of outcome, and particularly suggests the important role of patients' level of agency (or internal obstacles), positive and negative extra-therapeutic influences (especially interpersonal and work-related) and the patient-therapist/therapy relationship which evolves over time. The extent to which patient and therapy could find a congregation over the course of psychotherapy was important to facilitate change. The therapist was attributed an important role in this process. Moreover, patients in this study stressed that outcome or experienced changes not necessarily reflect therapy effects. As such, they nuance the role of psychotherapy and warn for simplistic linear process-outcome conclusions.

Based on these observations, we answer the third objective of this research project. First, our findings show that the JT method of clinical significance falls short to grasp the multi-dimensional nature of change as experienced by patients. This points at the limitations of a one-dimensional, symptom-based pre-post outcome evaluation. However, our findings moreover

question the classification of patients' outcome *as such* as well as their ability to differentiate patients' functioning after therapy in a clinically relevant way. Apart from a rough differentiation between patients who clearly improved ("recovered") and those who clearly worsened ("deteriorated"), the functional-dysfunctional normative comparison did not allow differentiating patients in a meaningful way with regard to their post-therapy functioning. Differences between patients who recovered and improved on the one hand and those who improved and did not improve on the other were less clear-cut. Specifically, the JT classifications did not grasp the more pessimistic experiences of those patients clustered under the header of "improved", nor the positive changes among those patients who "did not improve", hence omitting important clinically relevant nuances. Striking *within-sample heterogeneity* and c*ross-sample similarities* were observed, and as such resulted into fuzzy outcome classifications which appear more arbitrary than statistically intended.

Thus, whereas the JT method departs from an external norm and pre-defined outcome criterion aiming at universal applicability, our findings point at the necessity to evaluate patients' outcome by means of a first-person interpretation of the clinical significance. This is reflected in definitions of "personal recovery" and "positive health" (Delespaul et al., 2016; Slade & Longden, 2015). In contrast to external normative comparison, patients' outcome requires an internal level of comparison to enable conclusion on the clinical meaning of patient-specific change.

Hence, this dissertation arrives at the conclusion that the search for clinically useful outcome findings continues and should not reside in reliance on "the best available method" (so far) motivated by practical considerations. Stimulating this continuing endeavour, we argue for the adoption of a *broader* definition of clinical significance that expands a narrow statistical focus. First of all, this implies a diversification and complexification of research questions and applied methods and designs. Indicating explicitly *what* outcome, for *whom* or *what* purpose

has been evaluated should offer the right transparency on the applied methods and design to evaluate outcome. This, in a second step, then ensures drawing conclusions on the relevance of the findings for a specific research context. Ultimately, we should arrive at a foundation of nuanced and contextualized evidence that effectuates a clinically meaningful transition of findings into clinical care. Important implications for both policy and practice are discussed in light of these findings.

Nederlandstalige samenvatting

Het openen van de black box van psychotherapie-uitkomst: Een mixed methode studie naar het patiëntenperspectief

In dit proefschrift behandelen we de moeilijke relatie tussen academisch onderzoek en klinische praktijk waarbij we specifiek ingaan op de *klinische relevantie* van de bevindingen uit uitkomstenonderzoek. Daarbij adresseren we voornamelijk de implicaties voor RCT-gebaseerde bevindingen, gezien hun dominante positie als gouden standaard in het veld en hun daaruit voortvloeiende impact op de organisatie van de klinische praktijk. Dit onderzoek vertrekt van de veelgebruikte statistische definitie voor “klinische significantie”, zoals ze werd ontwikkeld door Jacobson en Truax (de “JT-methode”; 1991). Tot op vandaag vervult ze de positie van beste statistische indicatie voor de klinische relevantie van onderzoeksresultaten over de uitkomst van therapie. Op basis van een tweeledige formule (i.e., betrouwbare verandering pre-post therapie en evolutie van dysfunctionele naar functionele populatie) worden patiënten ingedeeld in vier mogelijk uitkomstcategorieën: “hersteld”, “verbeterd”, “niet veranderd/niet verbeterd” en “verslechterd”. De JT-methode biedt op die manier een patiënt-specifieke inschatting van de relevantie van de gemeten verandering en biedt op die manier inzage in de variatie binnenin algemene statistische uitkomstresultaten. Echter, ondanks deze belangrijke bijdrage aan de interpretate van statistische gemiddelde effecten, vertrekken we in

dit proefschrift van twee cruciale vragen die onbeantwoord blijven in de “black box” van therapie-uitkomst. Eerst en vooral: wat betekent klinisch significante verandering in relatie tot de ervaring van patiënten zelf? Ten tweede, welke veranderingsprocessen zijn onderliggend aan klinisch significante uitkomsten? De unidimensionale pre-post meting, ingebed in de JT-definitie laat niet toe hier antwoord op te bieden en vraagt daarom voor verdere conceptuele en methodologische ontwikkelingen. In de huidige studie benaderen we deze noodzaak door te vertrekken vanuit het perspectief van deze patiënten-participanten waarop uitkomstenonderzoek haar resultaten baseert en we presenteren daartoe we een representatief voorbeeld van het soort bevindingen die gehanteerd worden om de klinische praktijk te informeren.

Drie hoofddoelen leidden het onderzoek: 1) het onderzoeken van de ervaren veranderingen van patiënten om zo het begrip van “therapie-uitkomst” te verdiepen, 2) het in kaart brengen van de veranderingsprocessen die volgens patiënten betekenisvolle veranderingen hebben beïnvloed, om zo de mechanismen van verandering nader in kaart te brengen en tot slot, 3) het ondervragen van de mate waarin de statistische classificatie volgens Jacobson en Truax (1991) toelaat om klinisch betekenisvolle verandering in kaart te brengen volgens de ervaring van deze patiënten zelf. Het proefschrift schuift op die manier zowel een klinisch/conceptueel als methodologisch vraagstuk naar voren. Om een antwoord te bieden op de onderzoeksdoelen, wordt een mixed methode benadering vooropgesteld, bestaande uit drie lagen. Een sequentieel verklarend design zoals beschreven door Hesse-Biber (2010) wordt daarbij gevolgd. Concreet vertrekken we van de statistische (kwantitatieve) uitkomstevaluatie in termen van klinische significante (i.e., beschrijvende laag), waarna we deze uitbreiden middels een verdiepende kwalitatieve analyse van het perspectief van de patiënten op de ervaren veranderingen (i.e., exploratieve laag) en veranderingsprocessen (i.e., verklarende laag)

doorheen de therapie. Ultiem laat dit toe bestaande algemene statistische categorieën te openen en het perspectief van de patiënten in deze categorieën te bevragen.

Vier mixed methode studies werden uitgevoerd om antwoord te geven op deze centrale onderzoeksdoelen; in vijf opeenvolgende hoofdstukken worden ze weergegeven. Studie 1 (hoofdstuk II) presenteert een naturalistische studie in een residentiële setting in Vlaanderen (België). Binnen deze setting werd therapie-uitkomst op lange termijn bestudeerd, door kwantitatieve uitkomstenevaluatie te verrijken met een analyse van de ervaringen van voormalige patiënten, dit vijf of zes jaar na het beëindigen van de therapie. Een data-gedreven thematische analyse werd daartoe uitgevoerd op de cliëntveranderingsinterviews van 22 voormalige patiënten. Het resulterende tentatief conceptueel model van lange-termijn uitkomst wordt voorgesteld en besproken in hoofdstuk II. Deze studie vormt de pilootstudie voor het vervolg van het doctoraatsproject dat plaatsvindt in een gerandomiseerde gecontroleerde studie (RCT) naar de behandeling van depressie (Meganck et al., 2017). Studie 2 (zie hoofdstuk III en IV) onderzoekt daarbij het positieve uiteinde van het uitkomstencontinuüm. De patiënten uit de RCT met een "goede" pre-post therapie-uitkomst zoals gemeten met de Beck Depression Inventory (BDI), i.e. zij die "herstelden" en "verbeterden" in depressiesymptomen, werden geïncludeerd. De cliëntveranderingsinterviews van 28 herstelde en 19 verbeterde patiënten werd middels een grounded theory analyse onderzocht. Hoofdstuk III beschrijft de bevindingen met betrekken tot de ervaren veranderingen die deze patiënten ervoeren doorheen het therapieproces en vergelijkt het patiëntenperspectief op uitkomst in relatie tot de statistische classificaties van "hersteld" en "verbeterd". In hoofdstuk IV gaan we in op de processen van verandering en beïnvloedende factoren zoals beschreven door herstelde en verbeterde patiënten. Daarna, behandelen we deze vragen aan de negatieve kant van het uitkomstenspectrum in hoofdstuk V en VI. We bestuderen het perspectief van de patiënten in de RCT die een "onveranderde" (studie 3; hoofdstuk V) en "verslechterde" uitkomst vertoonden (gemeten met

de BDI). Een grounded theory analyse of de cliëntveranderingsinterviews resulteerde in een tentatief conceptueel model van "gebrek aan verandering" op basis van de ervaring van 19 patiënten (studie 3). Een interpretatieve fenomenologische analyse (IPA) op het narratief van drie verslechterde gevallen, laat in hoofdstuk VI toe de relaties tussen uitkomstenmeting en patiëntervaring nader te bestuderen. In alle hoofdstukken bespreken we de verhouding tussen de ervaringen van de patiënten enerzijds en de statistische classificatie anderzijds. In een laatste stap laat ons dit toe om uitspraken te doen over de mate waarin standaard uitkomstenonderzoek, en met name de veelgebruikte statistische indicatie van klinische relevante, een bruikbare vertaling naar de dagelijkse praktijk mogelijk maakt.

De algemene discussie (hoofdstuk VII) brengt een grondige integratie van de bevindingen uit de afzonderlijke studies, wat resulteert in een integratief conceptueel uitkomstenmodel. Als antwoord op het eerste doel van dit proefschrift, tonen de bevindingen dat uitkomst vanuit het perspectief van patiënten kan begrepen worden als een voortdurend proces, gekenmerkt door individuele verschillen. Patiënten beschrijven daarbij een holistische ervaring gekenmerkt door een multidimensionale aard van verandering. Verschillen tussen de patiënten wijzen daarbij op de mate van empowerment, internalisatie van inzichten en het bekomen van intra-en interpersoonlijke balans. Ook de mogelijkheid tot het vertalen van tools en inzichten uit de therapie in levensveranderingen of gedragsveranderingen en, of blijvende moeilijkheden gezien worden als normaal dan wel als hinderend en al van niet verdere therapeutsiche hulp vereisen, uit zich in verschillende mate over de patiënten heen.

Het antwoord op het tweede onderzoeksdoel, laat toe deze ervaringen en individuele verschillen nader te begrijpen en tentatief te verklaren. Eerst en vooral wijzen onze bevindingen op verschillende veranderingsprocessen – binnen en buiten de therapie – die interageren over tijd, wat in overeenstemming met de ervaren verandering een holistische therapie-ervaring toont. Therapie krijgt in de ervaring van de patiënt met andere woorden een plaats binnen diens

leven, persoon en context. Verschillen tussen patiënten wijzen daarbij voornamelijk op de rol van mate van agency (of interne obstakels) bij de patiënt, positieve en negatieve invloeden komende van buiten de therapiesessie (voornamelijk interpersoonlijk en jobgerelateerd) en de relatie tussen patiënt en therapeut/therapie die evolueert doorheen de tijd en het therapieproces. De mate waarin patiënt en therapie op elkaar geraken afgestemd doorheen het proces is daarbij belangrijk; de therapeut wordt daarbij een persoonlijke plaats toegewezen. Daarenboven geven de onderzochte patiënten aan dat de gemeten en ervaren uitkomst niet noodzakelijk een resultaat is van de effecten van de therapie. Op die manier waarschuwen ze dus voor simplistische, lineaire conclusies over de relatie tussen processen en uitkomst.

Op basis van bovenstaande observaties, beantwoorden we tot slot het derde onderzoeksdoel. We concluderen eerst en vooral dat de methode volgens Jacobson en Truax tekortschiet om de multidimensionale ervaring van patiënten te vatten. Dit wijst op de limitaties van een unidimensionale symptoomevaluatie als indicatie van therapeutische verandering. Echter, onze bevindingen stellen ook de uitkomstenclassificaties als dusdanig in vraag en daarbij hun bruikbaarheid om klinisch relevante conclusies te trekken over het functioneren van patiënten na therapie. Het meest duidelijke onderscheid werd geobserveerd tussen patiënten die duidelijk verbeterden (herstelden) en zij die verslechterden in symptomen. De normatieve vergelijking van patiënten ten aanzien van een functionele en dysfunctionele populatie liet daarnaast geen eenduidig onderscheid toe tussen patiënten die herstelden en verbeterden enerzijds en zij die verbeterden en niet verbeterden anderzijds. De JT-definitie slaagde er niet in de meer pessimistische ervaringen van “verbeterde” patiënten en de positieve ervaringen van “niet verbeterde” patiënten te vatten, met ongenuanceerde conclusies tot gevolg. Opvallende verschillen tussen patiënten *binnen een categorie* en belangrijke gelijkenissen tussen patiënten *over de categorieën heen*, maken van de classificaties vagere en meer arbitraire onderverdelingen dan waartoe ze bedoeld zijn. Waar de JT-methode dus vertrekt vanuit een

extern-bepaalde norm en vooraf bepaald criterium voor de evaluatie van uitkomst met als doel universele toepasbaarheid, tonen onze bevindingen noodzaak tot een andere benadering. Concreet sluiten onze bevindingen aan bij een eerste-persoonsdefinitie van uitkomst zoals vervat in definities van “persoonlijk herstel” en “positieve gezondheid” (Delespaul et al., 2016; Slade & Longden, 2015). De uitkomst van patiënten dient bij uitstek ten aanzien van een interne norm bepaald te worden, wil het klinisch van betekenis zijn als een patiënt-specifieke indicatie van verandering. De focus op symptomatische herstel wordt daarbij een nadruk op een gevoel van empowerment en agency, ongeacht blijvende symptomatische klachten.

Op die manier stellen we in dit proefschrift vast dat de zoektocht naar klinisch relevante uitkomstbevindingen verdergezet dient te worden en niet mag berusten in het gebruik van “de best beschikbare methode” (tot dusver), uit praktische overwegingen. Om deze verdere zoektocht te stimuleren, argumenteren we voor een *verruiming* van het begrip “klinische significantie”, voorbij een enge statistische interpretatie. Op de eerste plaats impliceert dit het stellen van meer diverse en complexe vragen met daarbij aansluitende variatie aan gehanteerde methoden en designs. Het expliciteren van *welke* uitkomst, *waarom*, voor *wie* of *welk doeleind* dient de nodige transparantie over de gehanteerde methode te bieden opdat het conclusies over de relevantie van de bevindingen voor de specifieke klinische praktijk toelaat. Ultiem dienen we te streven naar een fundament van genuanceerde en gecontextualiseerde evidentie die een klinisch bruikbare vertaling van onderzoeksbevindingen naar de praktijk toelaat. Verdere uitdagingen voor onderzoek, praktijk en beleidsvoering worden besproken in het licht hiervan.

Appendices

Appendix 1. Study 1. Clinical description, diagnoses and total scores (OQ-45) per patient.

Demogr.	Brief clinical description	Scores	
Female, 45-50	**Reason for consulting therapy centre:** Severe relational problems (described as psychological terror, domestic abuse by ex-husband) leading to isolation and eventually a suicide attempt. Entered inpatient therapy centre after two months in psychiatry. **Main difficulties:** Highly depending of others, clinging (describes herself as "a geisha"), e.g., initial idea behind treatment was becoming the ideal wife for her ex-husband. Highly insecure, low self-esteem. Suffering from a lack of energy, negative thoughts, struggling with femininity. **Diagnoses (DSM-IV):** Depressive disorder; PD NOS	**Start** **Post** **1yFU** **5yFU**	98 87 64 77
Female, 50-55	**Reason for consulting therapy centre:** Severe depression. Several losses (close relatives) preceding the depressive episode together with difficulties at work, leading to what she calls "the ultimate crash". Entered treatment after three months in psychiatry. **Main difficulties:** Life history of physical and verbal aggression (in family and relationships); single mother after difficult marriage. Central: always taken care of others (at home, at work, in intimate relationships). **Diagnoses (DSM-IV):** Depressive disorder; Dependant PD; Avoidant PD	**Start** **Post** **1yFU** **5yFU**	81 15 24 30
Male, 45-50	**Reason for consulting therapy centre:** "Heavy crash" following difficult period: losing job, father deceased, physical injury (cf. immobility), heavy work load and night shifts causing sleep deprivation. **Main difficulties:** History of problematic family situation caused by severe alcohol dependency father. Sexual abuse in childhood resulting in lack of trust in men and sex addiction. Extremely depending on others. **Diagnoses (DSM-IV):** Depressive disorder; Relational problems; Obsessive compulsive PD	**Start** **Post** **1yFU** **5yFU**	101 36 34 80
Female, 35-40	**Reason for consulting therapy centre:** Suffering from CVS since adolescence. Lived at home all her life, not able to take care of herself resulting in isolation; violent conflicts with parents. **Main difficulties:** CVS: severe level of inability, dependence and isolation. Perfectionism, inability in dealing with emotions. Unresolved past (e.g., victim of sexual harassment in childhood). **Diagnoses (DSM-IV):** Somatization disorder; PD NOS	**Start** **Post** **1yFU** **5yFU**	83 56 92
Male, 45-50	**Reason for consulting therapy centre:** "A crash", first physically (diagnosed with CVS), then mentally (depression). Verbal aggression towards wife and children (lead to seeking help). No experience with mental health care before. **Main difficulties:** History of violence in childhood (aggressive father). Perfectionism, inability in dealing with emotions (e.g., repressing anger and feelings in general), extreme fear of failure. **Diagnoses (DSM-IV):** Depressive disorder; Somatization disorder; Relational problems	**Start** **Post** **1yFU** **5yFU**	108 54 48

Male, 50-55	**Reason for consulting therapy centre:** Feelings of hopelessness and desperation. Outpatient care was experienced as insufficient. **Main difficulties:** A general feeling "not to be alive": existential crisis, experiencing life as meaningless. High demands and standards, not able to show vulnerability. History of "feeling different": as a child in his own family, later on in relation to friends and at work; history of rejection and exclusion. Suffering from fundamental loneliness. **Diagnoses (DSM-IV):** Depressive disorder (dysthymia); PD NOS	**Start** **Post** **1yFU** **5yFU**	60 45 86 75
Female, 55-60	**Reason for consulting therapy centre:** Feeling "stuck". Ambulant therapy sessions and short hospitalization in psychiatry were experienced as insufficient. **Main difficulties:** In general: difficult to live with self, interpersonal relations characterized by conflicts; difficulties in finding connection with others (ambivalence; distrust). Suffering from fundamental loneliness. **Diagnoses (DSM-IV):** Somatization disorder; Adjustment disorder; PD NOS	**Start** **Post** **1yFU** **5yFU**	102 81 86 89
Female, 35-40	**Reason for consulting therapy centre:** "Non-work-related burnout" (cf. extreme levels of stress and anxiety). Lived on "high speed" until "heavy crash" (e.g., social life, household became too much). Inpatient treatment necessary for being away from context. **Main difficulties:** History of psychological terror in family (aggressive father). Feelings of anxiety, high distress, easily overwhelmed. Low self-esteem and high need to please others. **Diagnoses (DSM-IV):** Anxiety disorder; PD NOS	**Start** **Post** **1yFU** **5yFU**	68 52 61 71
Female, 35-40	**Reason for consulting therapy centre:** Escalating relational problems; accumulating difficulties at work. Inpatient treatment necessary for being away from context. **Main difficulties:** Symbiotic relationship with mother (whom she describes as controlling and unable to give love). Therefore, not been able to develop own identity and personality apart from mother, leading to difficulties in other relationships (cf. highly depending on others). **Diagnoses (DSM-IV):** Adjustment disorder; PD NOS	**Start** **Post** **1yFU** **5yFU**	91 61 73 67
Female, 45-50	**Reason for consulting therapy centre:** Describes an overall feeling of "exhaustion" following family conflicts and marital problems. Break-up with life partner resulted in the loss of her stepson (and an important meaning in her life). Inpatient treatment necessary for being away from context. **Main difficulties:** Emotional vulnerability, no barriers to protect herself from influence and demands coming from others (extreme caring; over-empathic). Feeling different from others (and excluded from society). **Diagnoses (DSM-IV):** Depressive disorder; PD NOS	**Start** **Post** **1yFU** **5yFU**	93 58 93
Female, 45-50	**Reason for consulting therapy centre:** Severe depression after suicide attempt of husband and heavy care for children and household. Escalating relational problems (describes husband as manipulative and abusive); referred to inpatient therapy by relational therapist. **Main difficulties:** History of incest (childhood) and marital violence (psychological and physical). **Diagnoses (DSM-IV):** Depressive disorder; PTSD; Relational problems; PD NOS	**Start** **Post** **1yFU** **5yFU**	94 77 80 67
Female, 35-40	**Reason for consulting therapy centre:** Escalating problems in social and professional context (i.e., severe bullying and ending relationship), leading to increasing difficulties: inability to go to	**Start**	142

	work, isolation from friends, not able to do household, to take care of self. Referred to inpatient treatment by psychiatrist. **Main difficulties:** Severe bullying in youth; physical violence in family. Mother with severe alcohol dependency; rejected by father and grandparents. Describes a "fear of men" and fear of rejection. Alcohol abuse (self). **Diagnoses (DSM-IV):** Depressive disorder; PD NOS	**Post** **1yFU** **5yFU**	104 124 112
Female, 60-65	**Reason for consulting therapy centre:** Combination of occurrences in private life and work (high perfectionism, "work" was as a life mission; burnout). Inpatient treatment necessary for being away from context. **Main difficulties:** Struggles with the existential question on the meaning of life without children. **Diagnoses (DSM-IV):** Somatization disorder; Obsessive compulsive PD	**Start** **Post** **1yFU** **5yFU**	88 52 53 73
Female, 55-60	**Reason for consulting therapy centre:** Inpatient treatment after two years of sick leave due to burnout. Inpatient treatment necessary for being away from context (high demanding: husband and mother require daily mental and practical care/help). Feeling of despair, not being heard by previous therapist (treatment insufficient). **Main difficulties:** Extreme caring for others; low self-esteem; fundamental feeling of being different and abnormal. **Diagnoses (DSM-IV):** Depressive disorder; Obsessive compulsive PD	**Start** **Post** **1yFU** **5yFU**	100 49 56 68
Female, 45-50	**Reason for consulting therapy centre:** Many depressive episodes were treated outpatient before hospitalization. Suicidal thoughts lead to psychiatric hospitalization prior to inpatient therapy. **Main difficulties:** Problematic relationship with parents (psychological neglect); constant feelings of guilt; extreme caring for others. **Diagnoses (DSM-IV):** Depressive disorder; Somatization disorder; Avoidant PD; Dependent PD	**Start** **Post** **1yFU** **5yFU**	83 40 52 66
Female, 55-60	**Reason for consulting therapy centre:** "A breakdown" due to pressure at work. Depression, feeling of having "no way out". **Main difficulties:** Got separated from family at birth, reunited at age three. Has felt excluded by parents and siblings ever since. Suffered from dyslexia, which was only discovered lately. Low self-esteem as a consequence of abundant negative feedback on her cognitive abilities. Different person to the outside world. Difficulties in social situations (e.g., she cannot enter a public place alone or first). **Diagnoses (DSM-IV):** Depressive disorder; Avoidant PD; Dependent PD	**Start** **Post** **1yFU** **5yFU**	113 115 90 93
Female, 45-50	**Reason for consulting therapy centre:** Severe depression following escalating situation at home. Unable to work due to fibromyalgia (cf. isolation). Increasing feelings of loneliness and anxiety when alone at home and feelings of guilt towards family members (cf. not able to contribute to household). Inpatient treatment as a means of giving time, space and rest to all family members. **Main difficulties:** History of neglect as an infant (describes mother as cold, careless). Previous psychiatric treatment (as a teenager; two years of hospitalization) for anorexia nervosa. **Diagnoses (DSM-IV):** Depressive disorder (dysthymia); Relational problems; Dependent PD; Avoidant PD	**Start** **Post** **1yFU** **5yFU**	91 87 69 66

Female, 55-60	**Reason for consulting therapy centre:** General feeling of "not doing well" (regarding self, at work and in relationship). Referred to inpatient treatment by psychiatrist. **Main difficulties:** Child deceased due to a drug overdose; she feels guilty for his death. Violent marital relationship with ex-husband (psychological and physical aggression). Heavy depression and suicide attempt during that period. Autism (diagnosed after inpatient treatment) explains difficulties in relationships. **Diagnoses (DSM-IV):** Adjustment disorder; PD NOS	**Start** **Post** **1yFU** **5yFU**	101 73 66 95
Female, 60-65	**Reason for consulting therapy centre:** General anxiety and difficulties in relationship (e.g., dominant husband). **Main difficulties:** Father deceased when she was a child (she was not informed nor involved in the sickness process). Describes her mother as highly dominant (cf. breaking her down, inducing feelings of guilt). Anxiety (all her life, in relation to family, friends, partners: afraid not to be good enough, to do something wrong). **Diagnoses (DSM-IV):** Depressive disorder; PD NOS	**Start** **Post** **1yFU** **5yFU**	78 74 68 82
Female, 30-35	**Reason for consulting therapy centre**: After years of psychiatric treatment and individual therapy, referred to inpatient psychotherapy to address unresolved family issues and sexual abuse in childhood. **Main difficulties:** History of sexual abuse (childhood). Several hospitalizations in psychiatric care (three years in total) for the treatment of anorexia nervosa. Extreme fear of failure and perfectionism leading to depression and eating disorder (as a teenager). Difficult relationship with family following abuse and psychological difficulties. Negative body image. **Diagnoses (DSM-IV):** Adjustment disorder; PTSD; Borderline PD	**Start** **Post** **1yFU** **5yFU**	144 60 107 80
Female, 35-40	**Reason for consulting therapy centre:** Escalation of problems: accident (long revalidation); heavy job (working at night, no social life, isolation, sleep deprivation); sexual harassment (afraid to go out at night); unhealthy meals. Looked for help after aggressive outbursts at work. Reacted heavily (physically) on medication and was referred to inpatient treatment. **Main difficulties:** Describes family as harsh: speaking about difficulties and emotions is not done. Highly self-demanding (e.g., does not rest when ill). General distrust in others. **Diagnoses (DSM-IV):** Depressive disorder	**Start** **Post** **1yFU** **5yFU**	94 44 48
Female, 40-45	**Reason for consulting therapy centre:** "A crash" due to heavy care for children, all diagnosed with autism. Depression. **Main difficulties:** Difficulties due to (own) autism (diagnosed only years after inpatient treatment). Perfectionism; being easily overwhelmed. Difficulties in communicating feelings. Not easy to let others in (emotionally). **Diagnoses (DSM-IV):** Adjustment disorder; PD NOS	**Start** **Post** **1yFU** **5yFU**	73 102 116

Note n = 22. Information based on patients' own description of complaints, life history and reason for consulting the therapy centre as reported during the follow-up interview five to six years after ending therapy. Note that these brief clinical descriptions have the intention to provide more background information on patients (history, complaints other than symptomatic difficulties), but cannot be considered a thorough and complete anamnesis. Outcome scores: missing values were left blanc. DSM-IV diagnoses as assessed during the intake period, prior to inpatient treatment; information was provided by the centre's psychiatrist.

Appendix 2. Study 4. Case-specific findings: summary

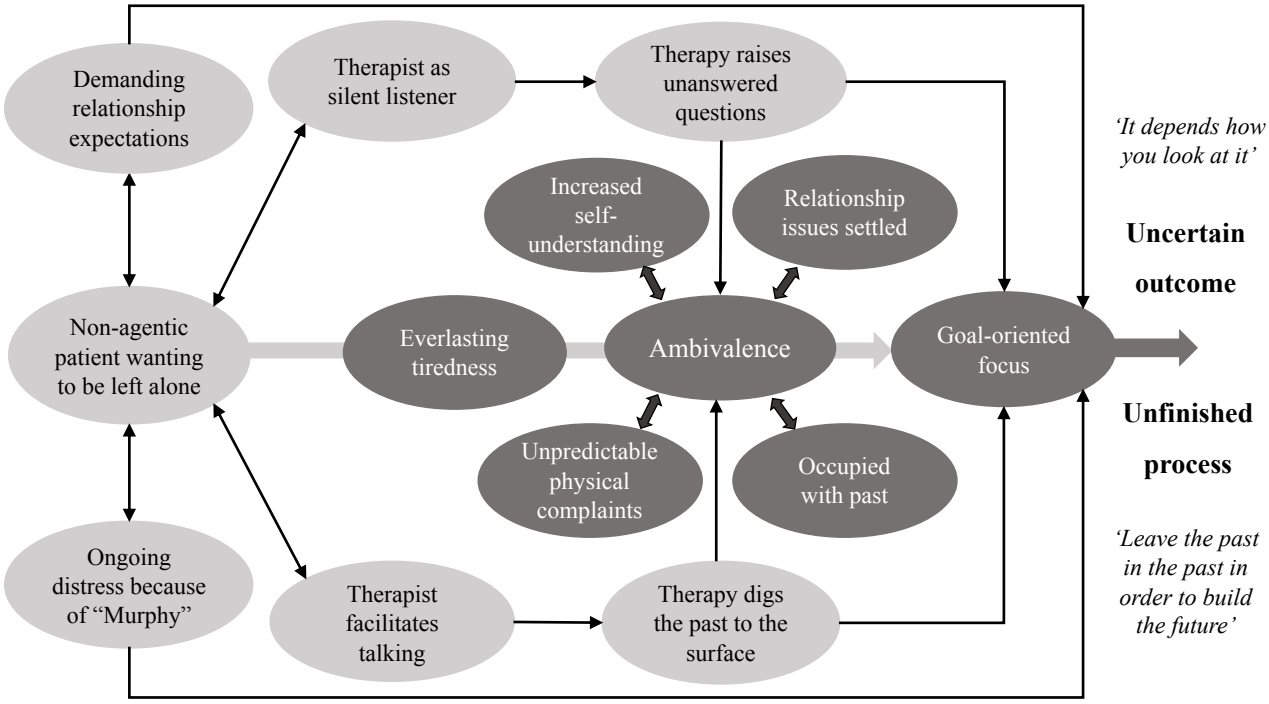
This appendix offers a summary of the core themes that were found for each individual case in the study. For each patient, a process model was designed to visualize the operating mechanisms.

Case I: Mark.

Taxonomy of Mark's central themes

Uncertain outcome following an unfinished process
Ambivalence: Is it good, bad, I don't know
Increased self-understanding; relationship issues settled; uncontrollable physical complaints; everlasting debilitating tiredness; preoccupied with past; goal-oriented focus
Therapy brings the past to the surface and raises unanswered questions
Therapist as silent listener who facilitates talking
Non-agentic patient wanting to be left alone
Ongoing distress because of "Murphy" and demanding relationship expectations*
*Murphy: professional misfortune experienced as "a pact with Murphy", e.g. ongoing lawsuit

Conceptual model of Mark's experiences of the therapy process and therapy outcome

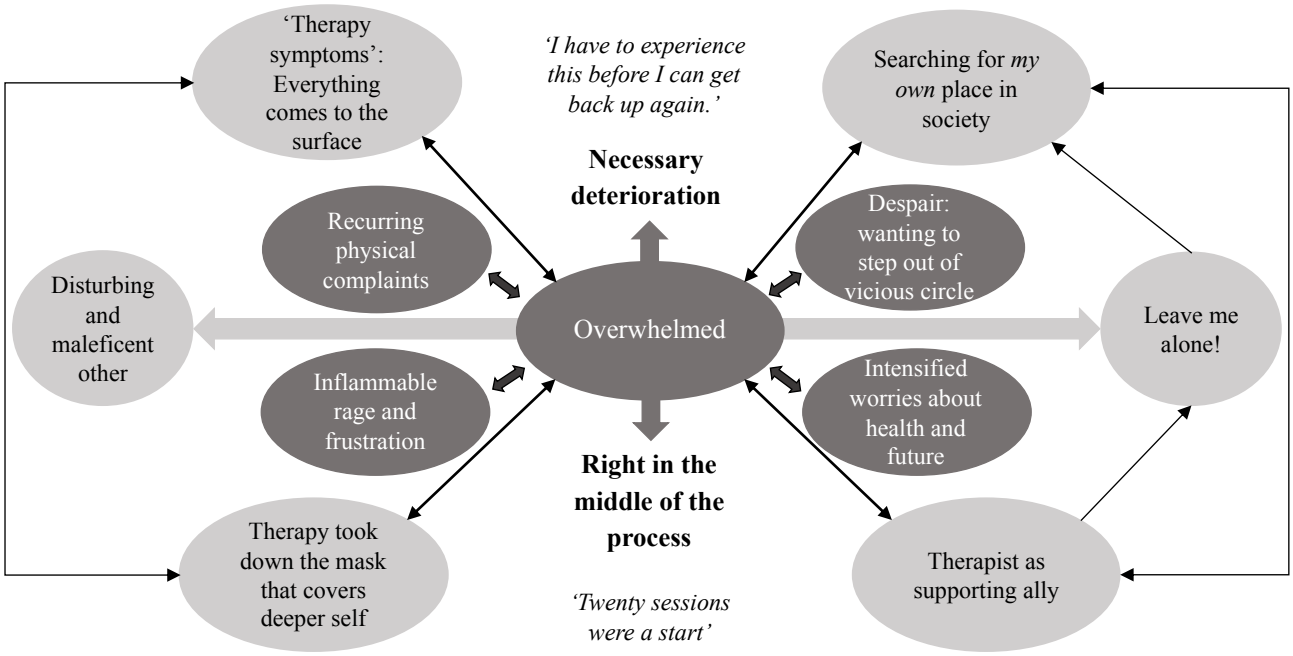


Case II: Ria.

Taxonomy of Ria's central themes

Necessary deterioration as a sign of being right in the middle of the process
Overwhelmed: Things pile up to a crisis *Recurring physical complaints; inflammable rage and frustration; intensified worries about health and future; despair: wanting to step out of the vicious circle* 'Therapy symptoms': everything comes to the surface as therapy took down the mask that covers deeper self Leave me alone! The disturbing and maleficent other (Ongoing distress: what will *my* future bring?) → Included in synthesis Finding my *own place* in society with the therapist as supporting ally

Conceptual model of Ria's experiences of the therapy process and outcome

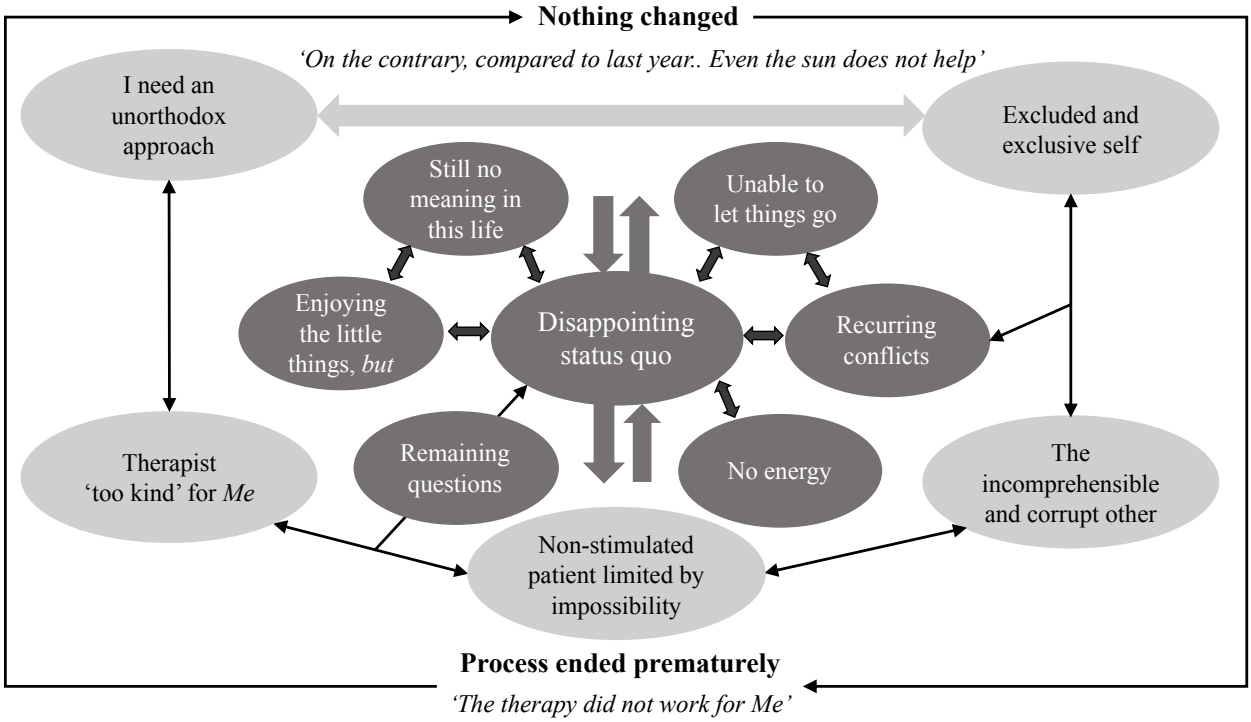


Case III: Miriam.

Taxonomy of Miriam's central themes

Nothing changed, therefore the process was ended prematurely
Disappointed with status quo: *No energy; unable to let things go; recurring conflicts; enjoying the little things but, still no meaning in life; remaining enigmas* Therapist too kind for Me, I need and unorthodox approach Non-stimulated patient limited by impossibility Excluded and exclusive self versus the incomprehensible and corrupt other (Ongoing necessity/reason to keep on going: 'my kids need me') → Included in synthesis

Conceptual model of Miriam's experiences of the therapy process and outcome



Appendix 3. Data Storage Fact Sheet

Name/identifier study: Opening the black box of psychotherapy outcome: A mixed methods study on patients' perspectives

Author: Melissa De Smet

Date: 26 June 2019

1. Contact details

==========

1a. Main researcher

- name: Melissa De Smet
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: Melissa.DeSmet@ugent.be

1b. Responsible Staff Member (ZAP)

- name: Reitske Meganck
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: Reitske.Meganck@ugent.be

If a response is not received when using the above contact details, please send an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies

==========

*** Reference of the publications in which the datasets are reported:**

Chapter II (Study 1):

De Smet, M. M., & Meganck, R. (2018). Understanding Long-term Outcome from the Patients' Perspective: A Mixed Methods Naturalistic Study on Inpatient Psychotherapy. *Psychologica Belgica, 58*(1), 276. doi: 10.5334/pb.432

Chapter III-VI (study 2-4):

De Smet, M. M., Meganck, R., Vannieuwenhove, K., Truijens, F. L., & Desmet, M. (2019). No Change? A Grounded Theory Analysis of Depressed Patients' Perspectives on Non-improvement in Psychotherapy. *Frontiers in Psychology, 10*(588). doi: 10.3389/fpsyg.2019.00588

De Smet, M. M., Meganck, R., De Geest, R., Norman, U. A., Truijens, F., & Desmet, M. (2019). What "Good Outcome" Means to Patients: Understanding Recovery and Improvement in Psychotherapy for Major Depression From a Mixed-Methods Perspective. *Journal of Counseling Psychology.* http://dx.doi.org/10.1037/cou0000362

De Smet, M. M., Meganck, R., Truijens, F., De Geest, R, Cornelis, S., Norman, U, & Desmet, M. (2019). Change Processes underlying "Good Outcome": A Mixed Methods Study on Recovered and Improved Patients' Experiences in Psychotherapy for Major Depression. *Psychotherapy Research* (manuscript accepted for publication pending revision)

De Smet, M.M., Cornelis, S., Acke, E., Reitske, M., Truijens, F., & Desmet, M. (2019). Making Sense of Deterioration in Psychotherapy: An Interpretative Phenomenological Analysis of Depressed Patients' Perspectives (manuscript forthcoming)

*** Which datasets in those publications does this sheet apply to?:**

Chapter II (Study 1):

Long-term outcome study conducted in a naturalistic inpatient setting. The internal process-outcome study was conducted between 2009 and 2012 ('Project van Integrale Kwaliteitszorg'; Rapport Outcome Questionnaire OQ-45.2 2009-2012; Deconinck, 2013). The long-term outcome study presented in this dissertation was led by Reitske Meganck.

Chapter III-VI (study 2-4):

The Ghent Psychotherapy Study, GPS, conducted by the Department of Psychoanalysis and Clinical Consulting (PP08), supervised by Mattias Desmet en Reitske Meganck (cf., Meganck et al., 2017, for the methodological protocol)

3. Information about the files that have been stored

==========

3a. Raw data

*** Have the raw data been stored by the main researcher? [x] YES / [ ] NO**

If NO, please justify:

*** On which platform are the raw data stored?**

- [x] research group file server
- [x] other (specify): patient-specific USB-sticks and ordners in locked archive

*** Who has direct access to the raw data (i.e., without intervention of another person)?**

- [x] main researcher
- [x] responsible ZAP
- [x] all members of the research group
- [x] other (specify): master thesis students, if participants gave informed consent for the use of data for educational purposes and after students have signed ethics clause

3b. Other files

*** Which other files have been stored?**

- [x] file(s) containing analyses. Specify: Nvivo codes and interview transcripts in doc files
- [x] files(s) containing information about informed consent
- [x] a file specifying legal and ethical provisions

*** On which platform are these other files stored?**

- [x] research group file server
- [x] other: patient-specific USB-sticks and ordners in locked archive

*** Who has direct access to these other files (i.e., without intervention of another person)?**

- [x] main researcher
- [x] responsible ZAP
- [x] all members of the research group

4. Reproduction

==========

*** Have the results been reproduced independently?: [x] NO**