

2012

# Does the depth of client experiencing predict good psychotherapy outcomes? A meta-analysis of treatment outcomes

Nikita Yeryomenko

Follow this and additional works at: <https://scholar.uwindsor.ca/etd>

---

## Recommended Citation

Yeryomenko, Nikita, "Does the depth of client experiencing predict good psychotherapy outcomes? A meta-analysis of treatment outcomes " (2012). *Electronic Theses and Dissertations*. 4847.  
<https://scholar.uwindsor.ca/etd/4847>

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email ([scholarship@uwindsor.ca](mailto:scholarship@uwindsor.ca)) or by telephone at 519-253-3000ext. 3208.

Does the depth of client experiencing predict good psychotherapy outcomes?

A meta-analysis of treatment outcomes.

By

Nikita Yeryomenko

A Thesis  
Submitted to the Faculty of Graduate Studies  
through Department of Psychology  
in Partial Fulfillment of the Requirements for  
the Degree of Master of Arts  
at the University of Windsor

Windsor, Ontario, Canada

2012

© 2012 Nikita Yeryomenko

Does the depth of client experiencing predict good psychotherapy outcomes?

A meta-analysis of treatment outcomes.

by

Nikita Yeryomenko

APPROVED BY:

---

M. Guarini, Department of Philosophy

---

D. Jackson, Department of Psychology

---

A. Pascual-Leone, Advisor

---

A. Scoboria, Chair of Defense

19 September 2012

## DECLARATION OF ORIGINALITY

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

## ABSTRACT

The Experiencing Scale (EXP), a measure of client's emotional processing, is often used in psychotherapy process research. While researchers agree that it predicts treatment outcomes, this relationship has not been systematically studied. This meta-analysis quantified the relationship between EXP and therapy outcomes using a total of 11 studies and 458 clients. Analysis indicated that peak EXP measured during the working phase was the strongest predictor of treatment outcomes,  $r = .236$ . Subgroup analyses indicated that working phase effects were moderated by the outcome measure modality. Early phase effects were moderated by the type of treatment and the treatment target. In accordance with the literature in the field, working phase EXP was found to be a significant predictor of clinical outcomes, although this relationship was influenced by a number of variables. Further research should look at the moderators between EXP and outcomes, and at processes that increase client experiencing.

## TABLE OF CONTENTS

DECLARATION OF ORIGINALITY .....	iii
ABSTRACT.....	iv
LIST OF TABLES .....	vii
LIST OF FIGURES .....	viii
LIST OF APPENDICES .....	ix
<b>CHAPTER 1. Introduction.....</b>	<b>1</b>
<b>Common Factors of Change.....</b>	<b>3</b>
<b>Client Experiencing as a Probable Common Factor.....</b>	<b>5</b>
<b>Rationale for the Current Study.....</b>	<b>9</b>
<b>Rationale for the Method.....</b>	<b>11</b>
<b>Current Study.....</b>	<b>12</b>
<b>Research Questions.....</b>	<b>15</b>
<b>CHAPTER 2. Methods .....</b>	<b>15</b>
<b>Literature Search.....</b>	<b>16</b>
<b>Coding.....</b>	<b>17</b>
<b>Data Analysis.....</b>	<b>18</b>
<b>Publication Bias.....</b>	<b>19</b>
<b>Choice of a Statistical Model.....</b>	<b>19</b>
<b>Weighting.....</b>	<b>19</b>
<b>Homogeneity of Effect Sizes.....</b>	<b>20</b>
<b>Hypothesis Testing: A Summary of Principal Analyses.....</b>	<b>21</b>
<b>CHAPTER 3. Results.....</b>	<b>21</b>
<b>Included Studies.....</b>	<b>21</b>
<b>Distribution of Effect Sizes Across Studies.....</b>	<b>24</b>
<b>Publication Bias .....</b>	<b>26</b>
<b>Individual Measures: EXP Process as a Predictor of Unique Treatment Outcomes.....</b>	<b>26</b>

<b>Total Index: EXP Process as an Overall Predictor of Treatment Outcome .....</b>	<b>29</b>
<b>Moderator Analysis.....</b>	<b>31</b>
<b>Clinical Judgments as a Moderator in EXP's Prediction of Outcome .....</b>	<b>31</b>
<b>Therapeutic Approach as a Moderator in EXP's Prediction of Outcome .....</b>	<b>33</b>
<b>Treatment Target as a Moderator in EXP's Prediction of Outcome .....</b>	<b>35</b>
<b>CHAPTER 4. Discussion .....</b>	<b>37</b>
<b>Implications for Research and Practice.....</b>	<b>40</b>
<b>Limitations.....</b>	<b>42</b>
<b>Future Research.....</b>	<b>43</b>
<b>REFERENCES .....</b>	<b>47</b>
<b>APPENDICES .....</b>	<b>56</b>
<b>VITA AUCTORIS .....</b>	<b>58</b>

## LIST OF TABLES

Table 1	
	<i>Study characteristics: Data sets and the reliability of EXP ratings.....</i> 23
Table 2	
	<i>Effect of EXP on individual outcome measures.....</i> 28
Table 3	
	<i>Total Index: Overall predictive power of EXP.....</i> 30
Table 4	
	<i>Effect of EXP based on whether Clinical Judgment outcome was used.....</i> 32
Table 5	
	<i>Effect of EXP based on type of treatment: Experiential vs. non-experiential therapy.....</i> 36
Table 6	
	<i>Treatment target as a moderator of EXP's prediction of outcome.....</i> 38



## LIST OF FIGURES

*Figure 1.*

The meta-analytic model.....14

*Figure 2.*

Stem-and-leaf plots of study effect sizes for each predictor measure.....25

*Figure 3.*

Treatment approaches and corresponding data sets.....34

LIST OF APPENDICES

APPENDIX A.....56  
APPENDIX B.....57

## CHAPTER 1

### Introduction

Some processes that occur during a psychotherapy session seem to hold promise for understanding how psychotherapy works. One such process is the depth with which clients experience their emerging feelings and thoughts (Greenberg & Pascual-Leone, 2006). Yet, no reviews have been conducted in the literature that might allow for an estimate of how important this particular process may be. In the hopes of addressing this question, the goal of the current study is to evaluate the role that a client's depth of experiencing plays in improving his or her outcome in psychotherapy. This introduction begins with a review of different types of psychotherapy research, outlining the major areas of outcome vs. process research. Next, I discuss the nature of process variables and common factors and explain their importance to understanding how psychotherapy works. Finally, I present a rationale for the current study, how it is embedded in the context of current psychotherapy research, and the intended impact of the study being proposed.

Over 30 years of studies and a multitude of meta-analyses have demonstrated the effect that psychotherapy interventions have on clients' treatment outcomes, and the evidence shows that therapy is superior to no treatment (Joyce, Wolfaardt, Sribney, & Aylwin, 2006; Budd & Hughes, 2009). Indeed, by several estimations the effect of psychotherapy is .8 standard deviations over no treatment (Lambert, 2005). While it is certain that therapy is better than no treatment, why it is effective is largely a mystery. Furthermore, a number of meta-analyses show evidence for the *Dodo Bird effect*: the finding that different psychotherapies are equally effective. At the same time, a few

studies also point to the superiority of certain intervention modalities for specific disorders and yet, differential processes are not usually identified (Joyce et al., 2006; Budd & Hughes, 2009). In any case, the Dodo Bird effect poses a significant problem for the field of psychotherapy. If seemingly disparate psychotherapies produce very similar outcomes, how does therapeutic change occur?

While studying the effectiveness of interventions and their comparative efficacy has led to the establishment of evidence-based practices, outcome research alone does not explain *how* therapy works or why outcomes are so similar (Garfield, 1990; Joyce et al., 2006). These questions, however, are the purview of “process research”. Finding processes through which therapy enacts change is doing a great service to empirically supported treatments; it adds the understanding that treatments work, but also which treatment components produce change, and which do not. It is possible then, through further research, to improve treatments and clarify them by removing any time-consuming procedures that do not produce change while emphasising those components that do.

Process research is an effort to explicate the causal processes of therapy that lead to change and addresses the shortcomings of outcome research (Joyce et al., 2006). Research and theory in this field has come to describe two general factors of change, treatment factors (i.e., interventions specific to a psychotherapy school such as insight or exposure) and common factors, which are ubiquitous to psychotherapy in general (i.e., client, therapist, the dyadic relationship, and general therapeutic process variables). There is an ongoing controversy in process research that revolves around these two kinds of factors and their relative contribution to outcomes (Nathan, 2004).

Rosenzweig (1936) was the first to propose the common factors idea that became increasingly popular in psychotherapy research, influencing psychologists such as Carl Rogers to elaborate his idea of the relationship as an overarching factor (Duncan, 2002). Common factors are also thought to be at least partly responsible for producing the Dodo Bird effect. Today, cumulative meta-analytic findings do indeed point towards the existence of common processes and mechanisms of change operating across therapies and which predict their eventual treatment outcome (Joyce et al., 2006; Budd & Hughes, 2009). Before proceeding with the overview of common factors of change, however, it is important to understand the relevant terminology and how it has evolved.

### **Common Factors of Change**

The strongest explanation of equal effectiveness among psychotherapies is the notion of factors that are empirically related to change but are not predicted or explained by theory underlying a given psychotherapeutic modality and yet, on closer examination, are present in most interventions (Oei & Shuttlewood, 1996). Three terms are usually used to describe such factors: placebo effect, nonspecific factors, and common factors.

Some have suggested that the *placebo effect* is an inappropriate term in a non-medical context, and that it also bears some negative connotations, thus downplaying the importance of some such processes of change (Oei & Shuttlewood, 1996). In medical context, a placebo is a psychological or psycho-physiological effect produced by the administration of an inert substance. In treatment research, such effects will be considered unwanted systematic variance by any researcher who looks for treatment-specific psychological mechanisms of change (Oei & Shuttlewood, 1996). Moreover, studies have shown that placebo control groups improve more than no-treatment groups

albeit less than treatment groups (Lambert, 2005), effectively demonstrating that the effect of psychotherapy reaches beyond that of a placebo alone.

Another formulation of variables of change underlying most psychotherapies is the notion of a nonspecific factor. This formulation is in contrast to the treatment specific (or technique) factors, and although *nonspecific factors* are conceptually more broad, the construct unfortunately tends to be reduced to no more than interpersonal and relationship factors (Castonguay & Holtforth, 2005). Consequentially, some studies have been criticized as placing relationship factors (such as therapeutic alliance) in the spotlight, while leaving out non-relationship factors responsible for change (Castonguay & Holtforth, 2005).

Finally, the truly *common factor* formulation has been outlined by Castonguay and Grosse Holtforth (2005). This formulation refers to those variables that operate across therapies, are not necessarily limited to the therapeutic alliance, such as facilitating new meaning-making and providing corrective experiences (Castonguay & Grosse Holtforth, 2005).

To place these issues in context one must consider the outcome literature at large, which indicates that therapeutic techniques account for no more than 10% of the variance in symptom change, while therapeutic alliance, the most well-researched and reliable measure of the relationship common factor, accounts for approximately the same amount (Castonguay & Beutler, 2006).

A recent APA Task Force has identified a number of treatment and common factors across dysphoric, anxiety, personality, and substance use disorders, as well as factors that only pertain to specific diagnostic categories (Beutler, Castonguay, & Follette, 2006;

Newman & Stiles, 2006; Critchfield & Benjamin, 2006; Haaga, McCrady, & Lebow, 2006). Therapeutic alliance, in particular, is perhaps the most robust and well-researched common factor. It can be conceptualized as an agreement between the client and the therapist on general goals of therapy, and specific tasks to accomplish those goals, as well as an emotional bond between the two individuals (Bordin, 1979). In a meta-analysis of 79 studies, the therapeutic alliance was found to be moderately but consistently related to outcome (Martin, Garske, & Davis, 2000). Therapeutic alliance is such a pervasive variable that it seems to eclipse other common factors and is used interchangeably with the broader term, as Castonguay and Grosse Holtforth (2005) pointed out.

While there is some progress towards identifying working process components of therapy (e.g., alliance), others have been largely overlooked. The depth of a client's experiencing, for example, is one variable that does not fit the participant characteristics identified in previous APA Task Forces on the study of common factors (Castonguay & Grosse Holtforth, 2005).

### **Client Experiencing as a Probable Common Factor**

Depth of experiencing refers to the emerging moment-by-moment integration of cognitions and affect, emotional arousal and emotional processing operating in tandem (Klein, Mathieu-Coughlan, & Kiesler, 1986). A deep level of experiencing is the process of making new meaning based on information derived from pre-verbal affective/meaning experiences that are attended to in the moment, what Gendlin (1981) also referred to as a *felt sense*. This in-the-moment process is different from simple emotional arousal (catharsis) or from purely cognitive processing of events in the absence of arousal. Operationalization of the experiencing construct is rooted in humanistic and experiential

approaches to psychotherapy, but as a new line of research studies has begun to suggest (e.g., Watson & Bedard, 2006), its application is not limited to these traditions. A similar concept is represented by *emotional engagement* or *emotional processing* in both the behavioural and cognitive approaches (Whelton, 2004). Over the recent years, this has also been receiving increasing attention in newer forms of psychodynamic psychotherapy (Whelton, 2004; Fosha, 2000).

While the initial process construct can seem somewhat abstract, a scale has been developed to assess a given depth of experiencing. The client Experiencing scale (EXP; Klein, Mathieu-Coughlan, & Kiesler, 1986) is a Likert-type, 7-point scale that uses trained raters to measure the depth of client's experiencing. To date, it is one of the most studied and validated measures of productive in-session process available in psychotherapy research, and it is often considered among scientist-practitioners to be a gold standard of measuring good experiential process (Greenberg & Pascual-Leone, 2006), although one would be hard-pressed to find an alternative measure that addresses the same construct.

Since the forays of psychotherapy process research, studies have used EXP in a number of different ways. Some have looked at EXP as a moment-by-moment outcome measure of good therapeutic process. In micro-process research of this kind, often referred to as process-to-process research, the focus is on measuring effects within a given session or moment. For example, Silberschatz, Fretter and Curtis (1986) investigated the effect of therapist's interpretations on client's productivity in the context of brief psychodynamic therapy; where client productivity was measured by a subsequent increase in EXP ratings (i.e., a positive impact on client's in-session work). Similarly,



Greenberg and Rice (1981) found that a gestalt *two-chair* intervention (i.e., enacting an imaginary dialogue with oneself) increased client's EXP, with implications for the client having a greater ability to cope with and resolve intrapersonal conflict. In another example, Macaulay, Toukmanian and Gordon (2007) found that when therapists showed greater empathy in emotion-focused psychotherapy, it also predicted higher levels of client's EXP. A number of studies from different schools of psychotherapy have used EXP as a measure of good in-session outcome on the theoretical assumptions that the process is important, and that the accumulation of these kinds of good moments actually yield good treatment outcomes later on. This leads to the empirical question: does EXP truly have practical importance?

Since the inception of this measure of experiential engagement, over a 100 studies internationally have used EXP in some capacity, as Hendricks (2009) pointed out in her broad review on this measure. While this is an impressive number, most of these studies are quite diverse in their purpose and research questions, such that most of them do not hope to use EXP as a predictor of post-treatment clinical outcomes. Nonetheless there is a small number of carefully designed studies that do specifically look at the relevance of EXP as a process variable predictive of final treatment outcome. Those studies are not limited to only humanistic or experiential psychotherapies (Greenberg & Pascual-Leone, 2006). For example, Goldman, Greenberg and Pos (2005) found that, in both client-centered and in emotion-focused psychotherapies, greater EXP during sessions was predictive of reduced symptoms of depression and increased self-esteem upon treatment completion. In a 12-week program of cognitive therapy, Castonguay, Goldfried, Wiser, Raue, and Hayes (1996) found that greater EXP during treatment was predictive of

improvement in depressive symptoms at the end of treatment. Similarly, in a comparative study of cognitive behavioural and emotion-focused therapies, Watson, McMullen, Prosser and Bedard (2011) again showed that, in both treatments, greater EXP was predictive of the following outcomes: decreased symptoms of depression, fewer interpersonal problems, less general psychological distress, as well as reduced dysfunctional attitudes and increased self-esteem. Interestingly, they also found that clients' affect regulation fully mediated the relationship between EXP and those outcomes.

The role of high EXP also seems important to treatments that emphasize interpersonal changes. Johnson and Greenberg (1988), in a study on emotionally focused therapy for couples, found that greater EXP for both partners was also characteristic of the most productive sessions of those couples who were most successful in therapy. Rudkin, Llewelyn, Hardy, Stiles and Barkham (2007) in their study that contrasted cognitive-behavioral therapy (CBT) and psychodynamic-interpersonal therapy (PI) found that PI therapy clients with good outcomes had higher levels of EXP than those that did not.

Finally, while there is clear evidence of the beneficial effects of psychotherapy on outcomes, the fact that gains are sometimes reversed or that clients may eventually relapse, still remains problematic (Ellison, Greenberg, Goldman, & Angus, 2009). It follows one of the aims of process research that that seeks to identify and facilitate good processes is because it may lead to an understanding of how to maintain treatment gains. So, establishing EXP as an index of a good psychotherapeutic process becomes important avenue of inquiry in this field and potentially for various ends.

Thus, research on process, outcome, and even some hypotheses about non-relapse have suggested that EXP may be a key construct with practical importance. Even so, despite this evidence, no *systematic* evaluation has been conducted to compile the evidence that EXP is a significant process variable as well as a probable common factor. So, while EXP is widely considered to be a gold standard and is believed to be predictive of outcomes, one would be hard pressed to find and cite a summative piece of research to that effect. One of the difficulties entailed in this task is that studies that have used EXP as a measure vary in their methodology and samples, among other characteristics, and EXP has been used in many ways across the literature. Summarizing the research findings to date on the general effect of EXP and its generalizability across treatment approaches would be helpful for the field by establishing client's depth of experiencing as an important process variable in psychotherapy.

### **Rationale for the Current Study**

The rationale of this study is in line with the general aims of process research: to identify what processes are related to, or cause, therapeutic changes to occur in the eventual service of good treatment outcomes. Finding process variables that go beyond the therapeutic alliance, especially if they generalize across therapeutic modalities, is helpful in informing the field of psychotherapy for both future research and practice. In the case of this particular study, there is a need to compile and examine the mounting evidence that EXP is an important predictor in its own right, as well as a probable common factor.

Showing that the depth of client experiencing is conclusively relevant across diverse treatment approaches is a key issue of inquiry because of its potential implications for

theory and intervention across different psychotherapy approaches. While interventions from a number of approaches provide explicit tasks such as evaluating maladaptive cognitions, deepening experience is a more implicit task that could transcend a host of other front-line interventions. While humanistic and experiential approaches have cultivated their intervention approaches to facilitate this, deepening experience is a matter of intention rather than a specific intervention per se, thus it is completed by the client and often facilitated through the *manner* in which an open-ended intervention is applied. Such a general intervention *strategy* may be applicable to most established treatments without interfering too much with their specific techniques, although some techniques would likely be more compatible with this additional implicit objective than others (e.g., interventions for facilitating insight, meaning-making, and awareness are highly compatible, while highly directed and behavioural tasks may not be). Having clear evidence about the impact of depth of client's experiencing on symptoms will be useful in developing new interventions and perhaps even in retrofitting established ones.

This proposed direction of research would also be useful, first of all, in shoring up the claims held by the growing body of existing process research that has used the EXP as an in-session outcome. At the same time it would be providing a foothold for future research of other key processes that may have been overlooked.

The first step in a research study on EXP is to compile and examine evidence supporting the notion of this variable as a probable common factor. There is some initial support for this idea (see Hendricks, 2009), and if a meta-analysis supports it, a promising research direction would be opened: the systematic study of which therapist processes or interventions lead to increasing client's EXP. In other words, the current

study could lead to practice-relevant implications for helping clinicians improve treatment outcomes for their clients. Furthermore, because process research usually relies on correlational designs (Joyce et al., 2006; Elliott, 2010), the current study might encourage experimental process research or help researchers answer Elliott's (2010) call for the use of more sophisticated statistical procedures such as structural equation modeling, which are naturally more demanding and complex. Without evidence from a systematic review showing the role of EXP, such studies on client process may never be conducted.

### **Rationale for the Method**

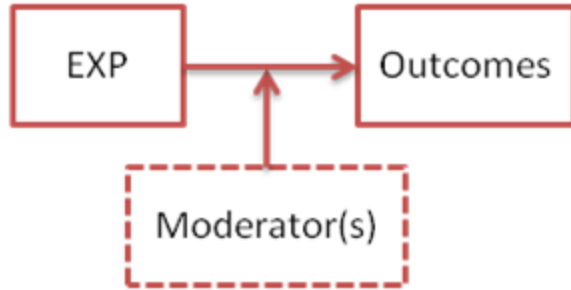
Hendricks (2009) has already provided an valuable narrative review on EXP and indicated that EXP was correlated with various psychotherapy outcome measures across studies. As indicated, other authors (e.g., Greenberg & Pascual-Leone, 2006) have further indicated that a small number of well designed process-to-outcome studies collectively suggest there may a broader effect of EXP that is waiting to be indexed. It follows then that given EXP is an important predictor of therapeutic outcome and regarded by many process researchers as a gold standard, the question of this study is: What is the actual general effect of deeper levels of experiencing? And, furthermore, does it generalize across studies with different characteristics? While a number of individual studies have been conducted in this area, replication in Psychology does not often lead to definitive results (Berkeljon & Baldwin, 2009). A meta-analysis, on the other hand, is a quantitative technique that allows one to address that problem by drawing upon data gathered from existing studies to answer research questions (Berkeljon & Baldwin, 2009). In doing so, a meta-analysis goes beyond a single sample, addressing the issue of research replication

and the generalization of findings across different populations. Because I hypothesize that study characteristics influence outcome, a meta-analytic approach provides additional research leverage (Berkeljon & Baldwin, 2009). With this approach it is possible to examine whether factors such as study design (experimental vs. quasi-experimental), clinical population or treatment type, or the method of outcome measurement have an effect on study findings. This breakdown of study-level covariates must be very selective and is limited by the number and size of study characteristics that can be coded.

### **Current Study**

This study relates client process to outcome. Greenberg (1986) postulated that outcomes in psychotherapy can be looked at three levels: immediate, intermediate, and ultimate outcomes. *Immediate outcomes* refer to a micro level change occurring within a single session, distinguishing good vs. poor therapeutic events (i.e., a productive moment in time or the success of a given intervention). At a slightly wider scope, *intermediate outcomes* refer to a broad pattern of change typically measured by session outcome measures, such as clinically relevant change in a client's attitudes and behaviours, distinguishing good vs. poor therapy sessions. Finally, *ultimate outcomes* are assessed at the end of therapy and at follow-ups. This scope of analysis paints the overall picture of clinical change and distinguishes a good vs. poor (i.e., successful vs. unsuccessful) course of psychotherapy. Both immediate and intermediate outcomes have been conceptualized as *small o's*, while the ultimate outcomes have come to be referred to as *large O's* (Greenberg & Pinsof, 1986). A complete description of the change process requires the study of all three levels of change and their relationships: how immediate outcomes influence intermediate outcomes, and later how these influence ultimate outcomes

(Greenberg, 1986). Previous research has studied EXP in two ways: as an outcome in itself (*small o*) brought about by other therapeutic processes (e.g., as therapist interventions or client insights), and as a process that predicts therapy outcomes (*large O*). To our knowledge no studies examined both roles of EXP simultaneously, although these can often be represented as different studies within a program of research. The current meta-analysis deals with the ultimate outcomes: how EXP is related to outcome measures at the end of therapy. This model is presented in Figure 1. Note that there is a possibility that the relationship between EXP and outcomes is moderated by other variables. In other words, contextual variables may change the relationship between EXP and outcome.



*Figure 1.* The meta-analytic model.



## **Research Questions**

This study is concerned with the following three questions:

1. Does EXP predict therapeutic outcomes? It is hypothesized that EXP positively predicts outcomes, such that as EXP increases, psychotherapy outcomes improve.
2. Is predictive power of EXP equal across therapeutic modalities? It is anticipated that this process variable successfully generalizes across different treatment approaches.
3. Are there variables that moderate the relationship between EXP and the clinical outcomes?

## **CHAPTER 2**

### **Methods**

This meta-analysis followed Berkeijon and Baldwin's (2009) outline of a number of steps for data synthesis, specifically with respect to psychotherapy outcome research. I will briefly present them, and then go on to explain how these steps are applied in the current study. (1) Research starts with a problem formulation, identical to how a conventional research study would proceed. This first step has been elaborated in the introduction of this document (above). The next step (2) is a literature search, given that "participants" in a meta-analysis are individual studies drawn from existing literature. Choosing studies with the right characteristics to address the research problem is essential. After the studies that qualify for a meta-analysis are gathered, (3) relevant data is extracted from them. Finally, (4) extracted data is subjected to statistical analysis and interpretation to answer the problems formulated at the beginning of the process.

## Literature Search

The literature search was conducted using PsycInfo, ERIC, Medline, and Social Sciences Abstracts databases. The following search terms were used: experiencing scale, emotional experiencing, client experiencing, patient experiencing, depth of experiencing, psychotherapy, psychotherapeutic processes, and EXP. However, because publication bias is an ever-present threat to the validity of meta-analysis (Berkeljon & Baldwin, 2009), we have contacted researchers to access additional unpublished material. This was done by making two open requests on well known research listserves with international memberships such as the listserv for the Society for Psychotherapy Research, the Society for the Exploration of Psychotherapy Integration, and the Short Term Dynamic Psychotherapy. Furthermore, direct inquiry was made to key researchers who were known to be involved in the use or study of the EXP. For the full list of contacted researchers see Appendix A.

Once the studies were collected, the following criteria were used to evaluate the suitability of each study for inclusion in the meta-analysis: (a) a given study must measure relevant constructs (both in-session EXP and clinical treatment outcomes), (b) it must feature a clinical population (i.e., participants with target clinical concerns such as Axis I and/or Axis II diagnosis), (c) it must report effect sizes or entail data that is convertible to effect sizes, of the relationship between EXP and treatment outcome, and finally (d) the included studies must have non-overlapping data sets. This last criterion is especially necessary because psychotherapy process research often relies on archival data from outcome studies; hence, one must take precautions to avoid the risk of redundancy in data. As it happened, a number of existing studies on the EXP have apparently drawn

from the same or similar archival databases (this is particularly the case at the York Psychotherapy Research Centre, Ontario Institute for Studies in Education (OISE), and University of Ottawa). Redundant data poses a very significant risk of biasing the meta-analytic results, because it violates the statistical assumption of independence of observations, which results in an overestimate of the mean effect and its precision (Wood, 2008). One of the corrections for this problem proposed in the methodological literature is to exclude the duplicates and only use the “main” article, the article with the largest data set (Wood, 2008). In the current study, research authors were personally contacted and consulted as needed regarding the potential overlap of their relevant data sets.

### **Coding**

Retrieved studies were coded independently by two researchers for the following information: (a) Treatment characteristics: treatment modality, treatment form, duration of treatment, intensity of treatment, therapist experience; (b) methodological characteristics: random assignment, blindness of EXP raters, presence of control groups, transcript sampling media, session EXP sampling method, within session EXP sampling method; (c) extrinsic characteristics: publication status, country of publication; (d) EXP characteristics: EXP type (mode, peak), EXP timing (early, working phase), percent of EXP data checked for reliability, hours of EXP rater training, EXP reliability ratings; (e) process measures and their reliabilities; (f) outcome measures; (g) effect sizes.

After all of the studies were coded by the primary researcher, the inter-rater agreement was assessed by the author and an undergraduate assistant. The assistant received two hours of training on the coding procedures on two of the 11 studies that

were used for the meta-analysis. Three of the studies were not coded for agreement because, following our request, the researchers sent the original SPSS data files, and two more were not checked for reliability because the effect sizes had to be computed from the raw data.

The second rater, the undergraduate assistant, conducted reliability check for 4 studies (36% reliability sample). Inter-rater agreement was  $r = .731$  for a total of 24 effect sizes. The only discrepancy appeared to be a missed negative sign for one of the original  $t$ -values, agreed to be a typo. Inter-rater agreement on timing of EXP (i.e., early or working phase), however, was low (kappa = .406), mostly due to difficulties in coding experiencing timings for clinical judgment studies that were not always clearly indicated. Agreement was perfect on treatment approach and treatment publication status. However, the rater was confused by the type of reliability coefficients the studies reported. While the numerical values matched, the types (i.e., ICC, Kappa) did not. A third rater, a senior researcher, coded treatment target reliability across all studies, resulting in excellent agreement (kappa = .783). All disagreements were resolved by consensus. Finally, clinical judgment emerged as an unexpected moderator during analysis, and therefore no inter-rater agreement rating was calculated for it. Clinical judgment codes were assigned to studies through a consensus with a senior researcher.

### **Data Analysis**

All analyses followed the methods outlined by Borenstein, Hedges, Higgins, and Rothstein (2009) and were computed in MS Excel 2007. Some analyses were re-checked with Comprehensive Meta-analysis (Version 2) software developed by Borenstein, Hedges, Higgins, and Rothstein (2005).

### **Publication Bias.**

Publication bias is an ever-present threat to a meta-analysis. While the methodology tried to address it by searching for unpublished material, it was important to evaluate whether such a bias is still present. In the current study, this was done through Orwin's Fail-safe  $N$  (Zakzanis, 2001). This statistic represents the number of hypothetical studies that, if added to already existing studies, will turn the result into a trivial effect size of  $r = .09$  or Cohen's  $d = 0.20$ , a convention proposed by Cohen (Zakzanis, 2001).

### **Choice of a Statistical Model.**

Meta-analyses can be conducted using either a fixed-effect or random-effects statistical model. A fixed-effect model assumes that the participant studies estimate the same true population effect size, with individual study differences being due to sampling error. A random-effects model assumes that different studies estimate different true population effects (Baker et al., 2009). Because a fixed-effect model does not take into account between study variability, which was expected due to clinical and methodological diversity (Thompson & Higgins, 2002), the random-effects model was more appropriate for the current study (see Baker et al., 2009). In addition, the goal of this meta-analysis was to generalize beyond the current sample, which a random-effects model is best suited for (Berkeljon & Baldwin, 2009).

### **Weighting.**

Because a meta-analysis pools effect sizes from multiple studies, each with different characteristics, it is important to weight those individual effect sizes relative to the contribution of the study from which each effect size came. This ensures, for example,

that smaller studies with large variance do not have an undue influence on the results of the analysis. Given the random-effects model, each effect size will be weighted by the inverse of the sum of within-study variance and between-study variance (Hedges & Vevea, 1998). A Monte Carlo simulation study has shown that when the random-effects assumption is true, the weighting method of Hedges and Vevea produces more precise estimates than an alternative of weighing by sample size (Marin-Martinez & Sanchez-Meca, 2009). If the random-effects assumption does not hold, however, the model is reduced to fixed-effects automatically (Borenstein et al., 2009).

### **Homogeneity of effect sizes.**

Diagnostics included a test for the homogeneity of effect sizes, or the assumption that effect sizes are dispersed around the mean to a degree no greater than what is expected due to sampling error alone. This is typically assessed by the  $Q$  test (Berkeljon & Baldwin, 2009). Statistical heterogeneity of effect sizes indicates that studies represent samples from different populations, rather than one uniform population (Baker et al., 2009), which might be expected given the diverse clinical and methodological nature of the studies in current research. Moreover, a  $Q$  test tends to be underpowered with small samples (Huedo-Medina, Sanchez-Meca, Marin-Martinez, & Botella, 2006), and even when it indicates a lack of heterogeneity, uneven patterns of variance can still be present (Hardy & Thompson, 1998). From what the practice of meta-analyses indicates, some heterogeneity will always be present, regardless of whether the  $Q$  test identifies its presence or not (Thompson & Higgins, 2002). Thus, a complementary statistic,  $I^2$  (Huedo-Medina et al., 2006; Sutton & Higgins, 2008) was also used to evaluate the extent of heterogeneity in the current study. Such heterogeneity may indicate an influence

of study-level characteristics (moderators) and is usually tested through a subsequent subgroup analysis.

### **Hypothesis Testing: A Summary of Principal Analyses**

The analyses were aimed at answering the following questions: (a) Does EXP predict therapeutic outcomes? (b) Does EXP have the same predictive power across therapeutic approaches? (c) Are there other moderators between EXP and therapeutic outcomes?

## **CHAPTER 3**

### **Results**

#### **Included Studies**

An initial literature search of published literature on the topic of the Experiencing Scale and psychotherapy outcomes resulted in 176 articles that could potentially qualify to be included in the meta-analysis. An additional 14 articles were obtained through listserv responses from individual researchers. Of the total 190 articles, 179 did not meet the selection criteria (see methods section) and also revealed a surprisingly high amount of redundancy in the field (see Appendix B for details). Of the studies that failed to meet the criteria, 109 did not use the Experiencing Scale. Forty two studies did examine the relationship between EXP and some outcome, that outcome was defined in terms of small “o’s” (Greenberg & Pinsof, 1986), while only studies that predicted final treatment outcome could meet the selection criteria. Eight more studies were excluded because they featured overlapping data sets (as Appendix B indicates, there were more than eight studies with overlapping data sets, but some of those were excluded because they studied small “o’s”). Six studies did not study clinical outcomes, another six studies did not have data convertible to effect size, four additional studies lacked vital information that was

requested but impossible to obtain (e.g., lost, destroyed), three studies were search duplicates, and one more study used EXP scale to validate another measure but did not link it to outcomes.

In the end, of the remaining 11 suitable studies used for this meta-analysis, 7 were published and 4 were unpublished dissertations/theses. From this set of 11, additional information was requested and received for 6 of the selected studies. Thus, a total of 458 clients from 11 studies were included in this meta-analysis. One of the studies (Hakim, 2010) compared two completely separate datasets and was therefore treated as two separate studies for the purposes of the analyses. Relevant study characteristics are presented in Table 1.

All of the 11 studies except one reported some form of inter-rater agreement rating for EXP data, and all but three indicated the amount of data that was subject to reliability check. Generally, at least 33% of data were checked (reliability samples ranging from 33% to 100% overlapping data). Cohen's Kappa and Intra-Class Correlation scores being in the good to excellent agreement range according to Fleiss (1981) and Fleiss (1986) benchmarks (e.g., kappa = .70 to .85).



Table 1

*Study characteristics: Data sets and the reliability of EXP ratings*

Study	N	Publication	Reliability of EXP ratings			Data checked
			Cohen's Kappa	Inter-rater $r$	Intra-Class Correlation	
Burgess (2012)	64	Unpublished	0.80			100%
Greenberg (1983)	28	Published				n/a
Hakim (2010; data set a)	28	Unpublished	0.70	0.74		33%
Hakim (2010; data set b)	29	Unpublished	0.70	0.80		33%
Makinen & Johnson (2006)	24	Published	0.83			n/a
Pachankis & Goldfried (2010)	52	Published			0.88	n/a
Pos et al. (2009)	73	Published	0.79			33%
Ralston (2006)	30	Unpublished	0.84	0.92		33%
Robichaud (2004)	37	Unpublished	0.85			33%
Rudkin et al. (2007)	8	Published			0.85	33%
Toukmanian et al. (2010)	19	Published		0.89		53%
Watson et al. (2011)	66	Published			0.83	69%

## **Distribution of Effect Sizes Across Studies**

While most studies conducted their ratings line-by-line (as dictated by the EXP manual), it has also become the convention to report findings in terms of either mode, peak (maximum EXP score), or both. The rationale for this in the literature is that *mode* is conventionally considered the best index of central tendency, while *peak* is considered an indicator of the upper limit of good process for a given therapy; and these are thought to be meaningful indices for clinical reasons (see Klien et al., 1986). Furthermore, many studies report the EXP scores at more than one point in the therapeutic process. Typically, EXP is measured early in therapy while the client-therapist rapport is being established (*early phase*), and later in treatment, when actual therapeutic work is being done (*working phase*).

For the purposes of this study early phase EXP was defined as measured during sessions one to four, and working phase during sessions five to two sessions before termination. This assignment of session numbers was made by browsing what the researchers reported as Early and Working phase sessions across the included studies. Mode scores represent the most consistently expressed level of experiencing, while the Peak scores represent the deepest level of experiencing, a moment of emotional “insight”. This meta-analysis used four predictors: early phase mode EXP, early phase peak EXP, working phase mode EXP, and working phase peak EXP Stem-and-leaf plots of the effect sizes by these categories are in Figure 2.

Early Mode EXP Fail-safe $N = -1$		Early Peak EXP Fail-safe $N = -1$		Working Mode EXP Fail-safe $N = 31$		Working Peak EXP Fail-safe $N = 57$	
Stem	Leaf	Stem	Leaf	Stem	Leaf	Stem	Leaf
-0.3		-0.3		-0.3		-0.3	
-0.2		-0.2		-0.2		-0.2	
-0.1	95	-0.1	59,54,45	-0.1		-0.1	
-0.0	66	-0.0	17	-0.0		-0.0	
0.0	00,64	0.0	32	0.0	65,68	0.0	05,31
0.1	20,42,77,88,88	0.1	64	0.1	00,70	0.1	21,79
0.2	00	0.2	27,30,46	0.2	68	0.2	
0.3		0.3	52	0.3		0.3	36,45
0.4		0.4		0.4		0.4	
0.5		0.5		0.5		0.5	
0.6		0.6		0.6	20	0.6	
0.7		0.7		0.7		0.7	16
0.8		0.8		0.8	21	0.8	
0.9		0.9		0.9		0.9	57

Figure 2. Stem-and-leaf plots of study effect sizes for each predictor measure.

## **Publication Bias**

This meta-analysis addressed publication bias by contacting researchers for unpublished studies, both directly and through listserves. Of the 11 studies that were included in the meta-analyses four were unpublished (36.4%). Additionally, Orwin's fail-safe  $N$  was computed for each of the four predictors and their corresponding effect sizes (see Figure 2). The results are typically judged by a benchmark proposed by Rosenthal (1979): if the fail-safe  $N$  is greater than the cut-off of  $5k + 10$  (where  $k$  is the number of studies) then such a number of unpublished studies is unlikely to exist.

For both early phase mode and peak effect sizes the fail-safe  $N$  was below zero. In terms of publication bias, this means that even if there are any unpublished studies with trivial effects they would not make any difference for the results if they were included in the analysis, because the observed effects in the early phase are already trivial.

In the working phase the fail-safe  $N$  that would be needed to trivialize the effects found for mode was a hypothetical sample size of 31 studies, and for peak, a sample size of 57. While fail-safe  $N$  for the effects of mode does not exceed its cut-off and peak fail-safe  $N$  does, given the specialized nature of the area and the effort put in obtaining the unpublished data, it is highly unlikely that such number of unpublished studies exists in reality. Thus, we feel that we have adequately addressed the publication bias.

## **Individual Measures: EXP Process as a Predictor of Unique Treatment Outcomes**

For exploratory purposes, separate meta-analyses were carried out for individual psychometric measures. The following outcome measures were commonly used and could be compared and compiled across process-outcome studies: Beck Depression Inventory (BDI), Inventory of Interpersonal Problems (IIP), Rosenberg Self-Esteem

Scale (RSE), Symptom Checklist-90-Revised (SCL-90-R), and a special category we labeled expert Clinical Judgment.

The “Clinical Judgment” category was assigned to an outcome measure when the outcome was decided based on some form of judgment by one or more expert clinicians (usually at some reported level of reliability) and this was done dichotomously, as in the resolution vs. non-resolution of a targeted personal or interpersonal problem (e.g., following gestalt, or couples therapy). After selecting a sample of participants with each of these types of outcomes, the researchers would then use a backwards prediction to compare them on their respective levels of process as measured by the EXP. For example, Burgess (2012) compared couples that achieved a “softening event” vs. the couples that did not, where the softening event occurred when the blaming partner approached their partner from a position of vulnerability and was understood and supported. Burgess then compared the two groups on their respective levels of EXP. Whether the couple achieved the softening event was based on an expert clinician’s judgment as well as blind ratings by independent raters on a checklist of in-session softening event markers. Burgess (2012) found that softened couples had higher levels of EXP during their best (working phase) sessions than couples who did not soften toward each other. The correlational nature of this type of design allowed us to determine to what degree higher levels of EXP predict a softening event.

The results of individual measure meta-analyses are summarized in Table 2, where  $k$  indicates the number of studies used for each analysis. There is no clear variation in the pattern that emerges from the meta-analyses of individual outcome measures. Significant and non-significant results occurred during both early and working phase for mode and

Table 2

*Effect of EXP on individual outcome measures*

Outcome	Predictor	<i>r</i> -effect				
		size	95% CI	<i>k</i>	<i>p</i>	<i>I</i> <sup>2</sup>
BDI	Early Mode	-0.133	-.270, .009	5	0.065	0.00
	Early Peak	-0.177	-.330, .016	6	0.032	21.70
	Working Mode	-0.256	-.437, .056	4	0.013	46.45
	Working Peak	-0.250	-.457, -.019	5	0.035	61.36
IIP	Early mode	-0.216	-.345, -.079	4	0.002	0.00
	Early peak	-0.281	-.404, -.147	4	0.000	0.00
	Working mode	-0.070	-.221, .084	3	0.372	0.00
	Working peak	-0.237	-.376, -.086	3	0.002	0.00
RSE	Early mode	0.117	-.138, .358	4	0.369	70.88
	Early peak	0.179	-.078, .414	4	0.171	71.36
	Working mode	0.183	.031, .327	3	0.019	0.00
	Working peak	0.179	-.168, .486	3	0.311	78.88
SCL-90-R	Early mode	-0.097	-.308, .123	7	0.387	71.74
	Early peak	-0.161	-.270, -.048	7	0.005	74.68
	Working mode	-0.141	-.305, .031	5	0.109	34.82
	Working peak	-0.266	-.425, -.092	5	0.003	40.12
Clinical	Early mode	0.137	-.074, .335	2	0.203	0.00
Judgment	Early peak	0.246	-	1	-	-
	Working mode	0.724	.457, .871	2	0.000	70.11
	Working peak	0.882	.383, .983	2	0.006	93.64

peak scores across the measures. Overall, the effect sizes had a generally consistent magnitude, with an exception of the clinical judgment measure that showed extremely large correlations with EXP (an issue that we will return to later).

As shown in table 2, a significant  $I^2$  heterogeneity among the effect sizes was present in some form for all of the measures, except for the Inventory of Interpersonal Problems (IIP). Normally, such heterogeneity would be further explored through moderator analysis. However, because of a small number of studies, moderator analysis for each of the individual measures was not a feasible strategy. The next step in analysis was the search for the overall predictive power of EXP. This was done by pooling the effect sizes within each study to compute a total index.

#### **Total Index: EXP Process as an Overall Predictor of Treatment Outcome**

The *total index* was obtained by taking the mean effect sizes across individual measures for each of the 11 suitable studies, and therefore represents a general clinical outcome. Because high scores on some measures indicate improvement in symptoms (e.g., Rosenberg Self Esteem), while high scores on others signal greater impairment (e.g., Beck Depression Inventory), the absolute value of correlation was used for the calculation of the total index. Because greater EXP is hypothesized to correlate with betterment of symptoms, when the actual correlation was not in the predicted direction, it was entered as a negative value and therefore detracted from the overall predictive power.

The results are shown in Table 3 for each phase of therapy. During the early phase, peak EXP was significantly associated with the total index,  $r = .124$ ,  $p = .048$ . The amount of dispersion of effect sizes was low,  $I^2 = 21.2\%$ , and did not exceed what would

Table 3

*Total Index: Overall predictive power of EXP*

Predictor	Mean <i>r</i> -effect size	95% CI	<i>k</i>	<i>p</i>	<i>I</i> <sup>2</sup>
Early Mode	0.093	-0.007, 0.191	10	0.0674	0.0
Early Peak	0.124	0.001, 0.243	10	0.0475	21.2
Working Mode	0.350	0.084, 0.567	7	0.0108	82.6
Working Peak	0.447	0.095, 0.670	8	0.0146	90.1



be expected by chance,  $Q(9) = 11.416, p = .248$ . Early in therapy, mode EXP was not found to be a statistically significant predictor of outcomes,  $r = 0.093, p = 0.067$ .

Working phase peak EXP was found to have the largest relationship with the total index,  $r = .447, p = .015$ , with a very high amount of effect size dispersion,  $I^2 = 90.1\%$ . It was followed by working phase mode EXP,  $r = .350, p = 0.011$ , with similarly high dispersion,  $I^2 = 82.6\%$ . In both instances the confidence intervals were very wide.

### **Moderator Analysis**

#### **Clinical Judgments as a Moderator in EXP's Prediction of Outcome.**

High heterogeneity for working phase mode and peak EXP effect sizes warranted further investigation. Visual inspection of effect sizes confirmed the suspicion that there often were extremely large effect sizes of EXP on clinical judgment outcomes during the working phase, which is believed to be the reason for the observed dispersion of effect sizes. To investigate further, subgroup analyses were done that compared those studies that used clinical judgment to those that did not (results are presented in Table 4). The general finding is, studies that used clinical judgment as an evaluation of treatment outcomes had a much higher average effect size than studies that did not. This difference was significant for both mode and peak EXP ( $Q(1) = 11.226, p < 0.001$ , and  $Q(1) = 5.126, p = 0.024$ , respectively). It is worthwhile to note that there were a total of three studies that used clinical judgment, and all were aligned with the experiential tradition (Burgess, 2012; Greenberg, 1983; Makinen & Johnson, 2006). Two of them were on experiential couples therapy.

Table 4

*Effect of EXP based on whether Clinical Judgment outcome was used*

Variable	Clinical		Mean $r$ - effect size	95% CI	$I^2$	$Q$	$p$
	Judgment as outcome ( $k$ )						
Working Mode	Not used (5)		0.157	0.023, 0.285	0.0	11.226	0.0008
	Used (2)		0.724	0.457, 0.871	70.1		
Working Peak	Not used (5)		0.236	0.110, 0.355	0.0	5.126	0.0236
	Used (2)		0.882	0.383, 0.983	93.6		

### **Therapeutic Approach as a Moderator in EXP's Prediction of Outcome.**

Subgroup analysis was used to establish whether the predictive power of EXP varies for different therapeutic approaches. Given that EXP is a construct that is derived from an experiential approach to psychotherapy, it seems particularly useful to examine whether predicative power of EXP on treatment outcomes is different across treatment approaches. Of the 11 process-outcome studies used in this meta-analysis, seven studies fit under the rubric of “experiential psychotherapy” and three represented multiple approaches that could be grouped under the heading “non-experiential therapy” (Hakim, 2010, compared two such treatments in the same study, IPT vs. CBT; Rudkin et al., 2007, compared CBT vs. psychodynamic-interpersonal therapy). Finally, one more study (Watson & Bedard, 2006) reported both experiential and non-experiential treatment comparisons in the same study (i.e., EFT vs. CBT). Thus, these studies yielded 13 independent data sets in total: eight on experiential therapy and five on non-experiential intervention approaches. Additionally, one of the studies (Watson et al., 2011) reported combined statistics for its EFT and CBT subgroups. Because of that, a related study that used a slightly smaller and overlapping dataset (Watson & Bedard, 2006) was substituted as it reported separate statistics for its EFT and CBT subgroups. This is the only instance when we used a study with an overlapping dataset (although in this case each study is used to uniquely address different research questions). See Figure 3 for details on specific treatment modalities. Given that clinical judgment outcomes have very large effect sizes and were only used by experiential studies, it was necessary to control for this confounding variable. Therefore, clinical judgment outcome studies were not used for

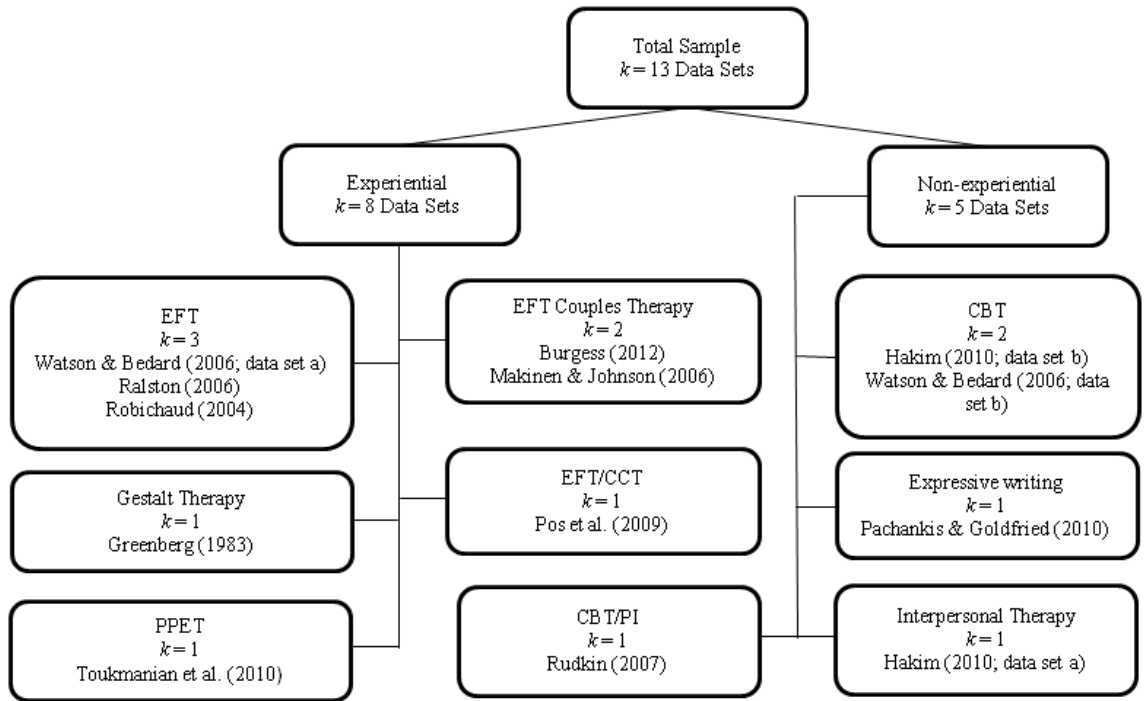


Figure 3. Treatment approaches and corresponding data sets.

this comparison, and all reported effects only reflect the relationship between process and symptom-based outcomes.

The predictive power of EXP was compared between experiential and non-experiential treatment studies with the  $Q$  heterogeneity test. The results are summarized in Table 5. The subgroup analysis indicated that peak EXP scores had a significantly greater predictive power for experiential studies as opposed to non-experiential studies (difference  $r = .278$ ) during the early phase of treatment ( $Q(1) = 5.527, p = 0.019$ ). Lack of other significant differences may reflect a small number of non-experiential studies and low statistical power.

#### **Treatment Target as a Moderator in EXP's Prediction of Outcome.**

Treatment studies can also be divided along the lines of target complains or the focus of treatment. While EXP has been discussed in the literature as a useful process for general therapeutic progress, it seems useful to conduct a subgroup analysis to examine whether the predictive power of EXP varies for different clusters of treatment concerns. Of the 13 data sets used in this meta-analysis (described above), six were specifically concerned with the treatment of “depression,” while seven were from studies that treated “interpersonal problems/trauma” (i.e., 2 treated complex relational trauma, 2 treated stressful interpersonal difficulties, 1 for prejudice-related trauma, and 2 used a couples modality of therapy). All three studies (including both couples therapies) that used clinical judgment as outcomes were also addressing “interpersonal problems/trauma”. To remove this confound, these three studies were excluded from the analyses and the examined relationships only reflected the prediction of symptom-based outcomes.

Table 5

*Effect of EXP based on type of treatment: Experiential vs. non-experiential therapy*

Variable	Treatment ( <i>k</i> )	Mean <i>r</i> -			
		effect size	95% CI	<i>Q</i>	<i>p</i>
Early Mode	Experiential (4)	.162	-.001, .317	1.618	0.203
	Non-experiential (5)	.005	-.174, .183		
Early Peak	Experiential (5)	.203	.053, .344	5.527	0.019
	Non-experiential (5)	-.075	-.247, .102		
Working Mode	Experiential (3)	0.206	0.025, 0.373	0.059	0.808
	Non-experiential (3)	0.170	-0.066, 0.388		
Working Peak	Experiential (4)	0.233	0.066, 0.388	0.335	0.563
	Non-experiential (3)	0.150	-0.086, 0.370		

In this way, the predictive power of EXP was compared between studies on individual therapies for depression vs. interpersonal problems/trauma with the  $Q$  heterogeneity test. The results are summarized in Table 6. The subgroup analysis indicated that only during the early phase of treatment, mode EXP was a better predictor of outcomes in studies that addressed depression as opposed to interpersonal problems/trauma ( $Q(1) = 4.212, p = 0.040$ ). It was impossible to make a comparison for working phase mode because of insufficient  $k$ . No other significant differences were found.

## CHAPTER 4

### **Discussion**

While the Experiencing Scale has often been referred to by process researchers as one of the best process measures to be predictive of outcomes, there has been no systematic review of it. In addition, a convention has been carried forward based on the clinical intuition of Klein et al. (1986) that EXP be reported using two different statistics: mode and peak. While researchers typically report both, sometimes they favor mode over peak in their analyses with the argument that a central tendency is more representative of what has happened in therapy. This however, should be an empirical question rather than one of rational theory, yet there has been no systematic evaluation of which EXP statistic is a better predictor of outcomes. To complicate matters further, EXP scores taken at the beginning of therapy may essentially represent different processes than those measured later in therapy, when the therapeutic alliance has been well established and some therapeutic work has been accomplished. As Pos et al. (2009) note, early EXP represents the baseline emotional processing capacity when clients first enter therapy, while

Table 6

*Treatment target as a moderator of EXP's prediction of outcome*

Variable	Target ( <i>k</i> )	Mean <i>r</i> -			
		effect size	95% CI	<i>Q</i>	<i>p</i>
Early Mode	Depression (6)	.192	.039, .336	4.212	0.040
	Interpersonal (3)	-.060	-0.242, 0.126		
Early Peak	Depression (6)	.136	-.036, .300	0.809	0.369
	Interpersonal (4)	.018	-.173, .207		
Working Mode*	Depression (5)	-	-	-	-
	Interpersonal (1)	-	-		
Working Peak	Depression (5)	.243	.090, .384	1.125	0.289
	Interpersonal (4)	.065	-.230, .349		

\*Insufficient *k* for the analysis.



working phase EXP represents the process of change, or deepening of emotional processing.

The results indicated that EXP measured during the early phase of therapy (sessions one to four) is not a strong predictor of ultimate therapy outcomes. When EXP was measured during the working phase of therapy (session five to two sessions before termination) EXP is indeed predictive of eventual outcomes, with peak EXP having a higher average effect size. Measuring outcome by way of clinical judgment (e.g., resolved vs. unresolved personal issues) was found to be an important moderator for the working phase predictors.

I briefly discuss this issue of the measurement moderator using peak EXP as an example, but the same conclusions apply to mode EXP findings as well. In short, the summary effect size that does not take this moderator into account ( $r = .447$ ) can be misleading. An extremely wide confidence interval indicates that there are really two different populations of data: that of the pen-and-paper psychometric measures (i.e., Beck Depression Inventory) which are predicted by EXP to some degree ( $r = .236$ ), and that of clinical judgment, an expert opinion, which shows EXP to have much more predictive power ( $r = .882$ ). This disparity cannot be easily reduced to researcher/judge bias given that clinical judgments on treatment outcome were always subject to ratings by separate judges (and corroborated by client self-report). An important caveat is that only one study out of three provided an index of reliability for clinical judgments and explicitly stated methodological precautions to ensure that such clinical judgments were also blind to treatment outcome.

Perhaps more critically, these approaches of measuring outcomes come from different assessment modalities and probably capture separate aspects of change that happens in therapy. It is possible that clinical judgment is a measure that is more sensitive to “true” change, or more of a personalized evaluation in a way that measures such as BDI are not. In fact, clinical judgment is the most "process-based" outcome measure. It is also the only evaluation of outcome that is based on actually viewing in-session performance to evaluate the treatment's success. This is a categorically different modality of assessment and may be more sensitive to the true or “lived” outcome, as clients grapple moment-by-moment with their troubles. That being stated, there were only two studies that measured working phase EXP and related it to clinical judgment outcome, so drawing any strong inferences about just a couple of studies is problematic. As such, relationship between the pen-and-paper measures and EXP is a more conservative and precise estimate of EXP’s predictive power.

### **Implications for Research and Practice**

Among known common factors, the therapeutic alliance is the most well-established predictor of in-session process on psychotherapy outcomes, with a correlation within .22 to .29 range (Horvath, 2005). This offers a point of comparison or context for our finding: Client’s peak EXP during the working phase of therapy as a predictor of post-treatment symptom reports has average effect in the range of .11 to .36 (95% CI), which is a correlation comparable to that of the working alliance. Thus, it is possible that client experiencing, as measured by the EXP, is another common factor of a similar magnitude and importance as the alliance, so further research on it will be a worthwhile investment for theory and practice of clinical psychology.

Furthermore, alliance and EXP clearly represent different constructs as demonstrated by a number of carefully controlled studies. For example, Ralston (2006) found correlations of  $r = -.003$  and  $r = .185$  between mode and peak EXP and alliance at pre-treatment. Similarly, Watson et al. (2011) found that correlations with the working phase alliance were  $r = .13$  during the early phase for both mode and peak EXP, and  $r = .15$  for mode and  $r = .29$  for peak EXP during the working phase. Finally, Pos et al. (2009) found that alliance correlates with EXP at  $r = -.05$  during the early phase and at  $r = .32$  during the working phase. These findings suggest that, at least in experiential therapies, there seems to be little to no relationship early in treatment, and then a small to moderate relationship eventually emerges over time.

In short, this suggests there is some relationship between these two different constructs, and they are not likely to be redundant with one another. Further to this point, the predictive power of EXP can be measured in the total absence of any therapeutic alliance as is demonstrated in the expressive writing study of Pachankis and Goldfried (2010). In psychotherapy research and theory, the alliance has been tentatively described as probably having a conditional relationship with deeper levels of experiencing, such that a strong alliance may provide a safe environment conducive to deeper levels of experiencing (Paivio & Pascual-Leone, 2010; Pos, Greenberg, Goldman, & Korman, 2003). Alliance was also considered within a mediational framework, such that during the working phase the relationship between therapeutic alliance and the outcomes was partially mediated by client experiencing (Pos et al., 2009). As such, the covariation of the alliance and other known process variables with EXP is also of interest. Furthermore, it is likely that the therapeutic alliance moderates the relationship between experiencing

and outcomes, such that experiencing predicts better outcomes only when therapeutic alliance is sufficiently strong.

### **Limitations**

These results should be considered while keeping in mind the limitations of correlational research and a very modest sample size. In case of significant correlations, it is possible that some other variable is fully or partially responsible for the relationship, such as therapeutic alliance, which we have discussed as a powerful common factor. Given a small number of studies it was impossible to control for this or other possible confounding factors. Another implication of having only 11 suitable outcome studies is that the summary effect sizes are estimated with less precision (i.e., the confidence intervals are wider) than would be the case if more studies were available. Finally, the statistical significance tests (such as *Q* tests) may be underpowered (Borenstein et al., 2009).

Another caveat to consider is lack of significant association between process and outcome, which occurs in the *early* phase mode EXP. The causal relationship may be there but possibly hidden by the highly interactive nature of psychotherapy. Stiles (1988) has described this phenomenon as the within-study variation in client requirements and corresponding therapist responsiveness. Essentially, in an effective therapy, a client with greater needs (e.g., a need for deeper experiencing, bodily-awareness, meaning-making) will be met with higher responsiveness from a psychotherapist using interventions and a therapeutic focus that is related to that need. So if a client needs constant validation of an emerging experience to get better (high requirement), a good psychotherapist will provide it frequently (high process). Similarly, if the client does not need constant validation to

get better (low requirement), a psychotherapist will move on to something more pressing, keeping the number of validations relatively low (low process). This responsiveness is clearly more sophisticated than the “more process is better” models of change. Unfortunately, however, limited factors and contexts further complicate the issue: For example, a client with exceptionally high requirements is less likely to have his or her requirements fully met, resulting in a poorer outcome, despite receiving a higher overall level of validations. Finally, client requirements and therapist responsiveness may vary within and between therapy sessions, such that the “true” relationship for each individual is actually part of a dynamic process in time. Thus, a group average of measured relationships between process and outcome may be diluted or even reversed, masking meaningful case-by-case relationships. In conclusion, it is possible that the true relationship between EXP and outcomes is actually higher than what the research designs and effect sizes in this meta-analysis have indicated.

### **Future Research**

This meta-analysis identified a number of challenges to current process research. A significant issue is that of overlapping datasets. Thirteen of the studies considered for this meta-analysis had samples that were shared with at least one other study (see Appendix B). When this occurred, a newer study with slightly different research questions would typically expand on the dataset used by the older one: either by (a) adding new participants or sampling methods (e.g., Pos et al., 2009, increases the sample size of Pos et al., 2003, which had also been used by Goldman et al., 2005) or (b) adding new comparison groups (e.g., Hakim, 2010, includes the same sample as Pos et al., 2003, and also compares these to rating of a different archival data set from Elkin et al., 1989).

While each individual study has made important contributions, this practice as a whole is problematic for a variety of reasons. First, the issue is often inadvertently obscured by the fact that an exact percent overlap among process studies using the same archival data sets is often undeclared and this could only be discerned in our study by personal communications with researchers. Second, in a study such as this if the overlapping datasets were mistakenly used they would violate the assumption of independence of observations, not only because the same people are used across the data sets, but also sometimes even the same EXP ratings and often the same outcome measures scores. Generally, researchers have acknowledged the redundancy of the data sets used in their studies; however, it would be desirable if more information were given, such as percent of overlapping data.

Although EXP has a magnitude of effect comparable to the therapeutic alliance, to be respected to the same degree it needs to have more support from studies that use a variety of treatments, outcome measures, and most importantly, different participants. With more studies of this kind an estimate of the predictive power of EXP can be defined much more precisely. Moreover, EXP is a measure of the “client experiencing” construct that can be measured by other means, such as the Client Emotional Productivity Scale developed by Greenberg, Auszra, and Herrmann (2007). The construct as a whole deserves more attention from individual studies, and as suggested earlier, would benefit from research designs incorporating multiple measures of client experiencing in a structural equation framework. A similar suggestion concerns the outcome measures. Given the results of this meta-analysis, EXP has different relationship with different modalities of outcomes. It is possible that some effects may be lost because only paper-and-pencil measures are

included, while others may perhaps be overstated because of the use of expert opinion as an outcome. Incorporating multiple outcome modalities as well as multivariate analyses would be a welcome addition to the field.

While this study has produced important findings on the role of EXP process-to-outcome (final treatment outcome), the original proposal aspired to also examine the role of other client and therapist in-session processes as predictors of EXP. This extension has not been possible given the time limits of this project and the intensive nature of coding. There are an estimated 32 studies that would potentially qualify for the process-to-process meta-analysis. However, as stated earlier, significant predictive power of EXP justifies such process-to-process study. Given that EXP correlates with final treatment outcomes, it would be useful for practitioners to know how exactly they can increase client experiencing while working in-session. Therapists do not have control over the baseline client experiencing levels, and the effects of early experiencing on final outcomes were found to be trivial to small. That is not the case with working phase client experiencing. As stated earlier, experiencing in the working phase represents a process of change from the baseline levels. Assuming therapists can influence this change by some intervention strategies (therapist processes), such as empathic responding or validation, they can shape the outcomes. If the process-to-process study were to be done, its focus should be on working phase EXP, particularly the peak scores, as they show a bigger effect on outcome. The results of such study would complement this meta-analysis in informing practitioners what they can do to help their clients beyond the specific interventions of any given therapeutic approach.

Finally, given that EXP is a potential common factor, its predictive power should be investigated across therapeutic modalities and controlled for potential confounds. More specifically, it can be beneficial to see how therapeutic alliance and EXP interact with each other in predicting the outcomes.



## REFERENCES

- Baker, W. L., White, C. M., Cappelleri, J. C., Kluger, J., & Coleman, C. I. (2009). Understanding heterogeneity in meta-analysis: The role of meta-regression. *International Journal of Clinical Practice, 63*(10), 1426-1434.
- Berkeljon, A., & Baldwin, S. A. (2009). An introduction to meta-analysis for psychotherapy outcome research. *Psychotherapy Research, 19*(4-5), 511-518.
- Beutler, L. E., Castonguay, L. G., & Follette, W. C. (2006). Therapeutic factors in dysphoric disorders. *Journal of Clinical Psychology, 62*(6), 639-647.
- Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research & Practice, 16*(3), 252-260.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to Meta-analysis*. Chichester, UK: Wiley.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2005). Comprehensive Meta-analysis (Version 2) [computer software]. Englewood, NY: Biostat.
- Budd, R., & Hughes, I. (2009). The Dodo Bird verdict — controversial, inevitable and important: A Commentary on 30 years of meta-analyses. *Clinical Psychology and Psychotherapy, 16*, 510-522.
- \*Burgess Moser, M. (2012). *The cognitive-affective and behavioural impact of emotionally focused couple therapy*. (Unpublished doctoral dissertation). University of Ottawa, Ottawa.

- Castonguay, L. G., & Beutler, L. E. (2006). Principles of therapeutic change: A Task Force on participants, relationships, and technique factors. *Journal of Clinical Psychology, 62*(6), 631-638.
- Castonguay, L. G., & Grosse Holtforth, M. (2005). Change is psychotherapy: A plea for no more “nonspecific” and false dichotomies. *Clinical Psychology: Science and Practice, 12*(2), 198-201.
- Castonguay, L. G., Goldfried, M. R., Wisner, S., Raue, P. J., & Hayes, A. M. (1996). Predicting the effect of cognitive therapy for depression: A study of unique and common factors. *Journal of Consulting and Clinical Psychology, 64*(3), 497-504.
- Cohen, J. (1994). A power primer. *Psychological Bulletin, 112*(1), 155-159.
- Critchfield, K. L., & Benjamin, L. S. (2006). Principles for psychosocial treatment of personality disorder: Summary of the APA Division 12 Task Force / NASPR review. *Journal of Clinical Psychology, 62*(6), 661-674.
- Duncan, B. (2002). The founder of common factors: A conversation with Saul Rosenzweig. *Journal of Psychotherapy Integration, 12*, 10-31.
- Elkin, I., Shea, M. T., Watkins, J. T., Imber, S. D., Sotsky, S. M., Collins, J. F., Glass, D. R., Pilkonis, P. A., Leber, W. R., & Docherty, J. P. (1989). National Institute of Mental Health treatment of depression collaborative research program: General effectiveness of treatments. *Archives of General Psychiatry, 46*(11), 971-982.
- Elliott, R. (2010). Psychotherapy change process research: Realizing the promise. *Psychotherapy Research, 20*(2), 123-135.

- Ellison, J. A., Greenberg, L. S., Goldman, R. N., & Angus, L. (2009). Maintenance of gains following experiential therapies for depression. *Journal of Consulting and Clinical Psychology, 77*(1), 103-112.
- Fleiss, J. L. (1981). *Statistical rates for measures and proportions* (2nd ed.). New York: Wiley.
- Fleiss, J. L. (1986). *The design and analysis of clinical experiments*. New York: Wiley.
- Fosha, D. (2000). *The Transforming Power of Affect*. New York: Basic Books.
- Garfield, S. L. (1990). Issues and methods in psychotherapy process research. *Journal of Consulting and Clinical Psychology, 58*(3), 273-280.
- Gendlin, E. T. (1981). *Focusing*. Toronto: Bantam Books.
- Goldman, R. N., Greenberg, L. S., & Pos, A. E. (2005). Depth of emotional experience and outcome. *Psychotherapy Research, 15*(3), 248-260.
- Greenberg, L. S. & Pascual-Leone, A. (2006). Emotion in Psychotherapy: A practice-friendly research review. *Journal of Clinical Psychology: In Session, 62*, 611-630.
- Greenberg, L. S. & Pinsof, W. M. (1986). Process research: Current trends and future perspectives. In L. S. Greenberg & W. M. Pinsof (Eds.), *The psychotherapeutic process: A research handbook*. New York: Guilford.
- Greenberg, L. S., & Rice, L. N. (1981). The specific effects of a gestalt intervention. *Psychotherapy: Theory, Research and Practice, 18*(1), 31-37.
- \*Greenberg, L. S. (1983). Toward a task analysis of conflict resolution in gestalt therapy. *Psychotherapy: Theory, Research and Practice, 20*(2), 190-201.
- Greenberg, L. S. (1986). Change process research. *Journal of Consulting and Clinical Psychology, 54*(1), 4-9.

- Greenberg, L. S., & Pascual-Leone, A. (2006). Emotion in psychotherapy: A practice-friendly research review. *Journal of Clinical Psychology: In Session*, 62(5), 611-630.
- Greenberg, L. S., Ausra, L., & Hermann, I. R. (2007). The relationship among emotional productivity, emotional arousal and outcome in experiential therapy of depression. *Psychotherapy Research*, 17(4), 482-493.
- Haaga, D. A. F., McCrady, B., & Lebow, J. (2006). Integrative principles for treating substance use disorders. *Journal of Clinical Psychology*, 62(6), 675-684.
- \*Hakim, L. Z. (2010). *A comparative analysis of the qualities of therapists' communications in three empirically-supported psychotherapies*. (Unpublished doctoral dissertation). York University, Toronto, ON.
- Hardy, R. J., & Thompson, S. G. (1998). Detecting and describing heterogeneity in meta-analysis. *Statistics in Medicine*, 17, 841-856.
- Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods*, 3(4), 486-504.
- Hendricks, M. (2009). Experiencing level: An instance of developing a variable from a first person process so it can be reliably measured and taught. *Journal of Consciousness Studies*, 16(10-12), 129-155.
- Horvath, A. O. (2005). The therapeutic relationship: Research and theory. An introduction to the special issue. *Psychotherapy Research*, 15(1-2), 3-7.
- Huedo-Medina, T. B., Sanchez-Meca, J., Marin-Martinez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or  $I^2$  index? *Psychological Methods*, 11(2), 193-206.

- Jackson, S. (2011). *Change processes in emotion-focused and interpersonal psychotherapies: A comparative study*. (Unpublished doctoral dissertation). York University, Toronto, ON.
- Johnson, S. M., & Greenberg, L. S. (1988). Relating process to outcome in marital therapy. *Journal of Marital and Family Therapy, 14*(2), 175-183.
- Joyce, A. S., Wolfaardt, U., Sribney, C., & Aylwin, S. (2006). Psychotherapy research at the start of the 21st century: The persistence of the art versus science controversy. *Canadian Journal of Psychiatry, 51*(13), 797-809.
- Klein, M. H., Mathieu-Coughlan, P., & Kiesler, D. J. (1986). The experiencing scales. In L. S. Greenberg & W. M. Pinsof (Eds.), *The psychotherapeutic process: A research handbook* (pp. 21-71). New York, NY: Guilford Press.
- Lambert, M. J. (2005). Early response in psychotherapy: Further evidence for the importance of common factors rather than “placebo effects”. *Journal of Clinical Psychology, 61*(7), 855-869.
- Levitt, H., Korman, Y., & Angus, L. (2000). A metaphor analysis in treatments of depression: metaphor as a marker of change. *Counselling Psychology Quarterly, 13*(1), 23-35.
- Littell, J. H., Corcoran, J., & Pillai, V. K. (2008). *Systematic reviews and meta-analysis*. Oxford, U.K.: Oxford University Press.
- Macaulay, H. L., Toukmanian, S. G., & Gordon, K. M. (2007). Attunement as the core of therapist-expressed empathy. *Canadian Journal of Counselling, 41*(4), 244-255.
- Makinen, J. A. (2005). *Resolving attachment injuries in couples: Relating process to outcome*. (Unpublished doctoral dissertation). University of Ottawa, Ottawa, ON.

- \*Makinen, J. A., & Johnson, S. M. (2006). Resolving attachment injuries in couples using emotionally focused therapy: Steps towards forgiveness and reconciliation. *Journal of Consulting and Clinical Psychology, 74*(6), 1055-1064.
- Marin-Martinez, F., & Sanchez-Meca, J. (2009). Weighting by inverse variance or by sample size in random effects meta-analysis. *Educational and Psychological Measurement, 70*(1), 56-73.
- Martin, D. J., Garske, J. P., & Davis, M. K. (2000). Relation of the Therapeutic Alliance with outcome and other variables: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 68*(3), 438-450.
- Missirilan, T. M. (2011). *A comparative study of the nature of change processes in emotion focused and cognitive-behavioural psychotherapies for depression*. (Unpublished doctoral dissertation). York University, Toronto, ON.
- Naaman, S., Pappas, J. D., Makinen, J., Zuccarini, D., & Johnson-Douglas, S. (2005). Treating attachment injured couples with emotionally focused therapy: A case study approach. *Psychiatry, 68*(1), 55-77.
- Nathan, P. E. (2004). The evidence base for evidence-based mental health treatments: Four continuing controversies. *Brief Treatment and Crisis Intervention, 4*(3), 243-254.
- Newman, M. G., & Stiles, W. B. (2006). Therapeutic factors in treating anxiety disorders. *Journal of Clinical Psychology, 62*(6), 649-659.
- Oei, T. P. S., & Shuttlewood, G. J. (1996). Specific and nonspecific factors in psychotherapy: A case of cognitive therapy for depression. *Clinical Psychology Review, 16*(2), 83-103.

- \*Pachankis, J. E., & Goldfried, M. R. (2010). Expressive writing for gay-related stress: Psychosocial benefits and mechanisms underlying improvement. *Journal of Consulting and Clinical Psychology, 78*(1), 98-110.
- Pascual-Leone, A. (2009). Dynamic emotional processing in experiential therapy: Two steps forward, one step back. *Journal of Consulting and Clinical Psychology, 77*(1), 113-126.
- Pascual-Leone, A., & Greenberg, L. S. (2007). Emotional processing in experiential therapy: Why "the only way out is through." *Journal of Consulting and Clinical Psychology, 75*(6), 875-887.
- Pos, A. E., & Escobar, M. (2011, May). *Early therapy processes differences and outcome for experiential therapy for depression: Process subtypes in experiential treatment of depression*. Paper presented at the annual meeting of Society for the Exploration of Psychotherapy Integration, Washington, DC.
- \*Pos, A. E., Greenberg, L. S., & Warwar, S. H. (2009). Testing a model of change in the experiential treatment of depression. *Journal of Consulting and Clinical Psychology, 77*(6), 1055-1066.
- Pos, A. E., Greenberg, L. S., Goldman, R. N., & Korman, L. M. (2003). Emotional processing during experiential treatment of depression. *Journal of Consulting and Clinical Psychology, 71*(6), 1007-1016.
- \*Ralston, M. B. (2006). *Imaginal confrontation versus evocative empathy in emotion focused trauma therapy*. (Unpublished doctoral dissertation). University of Windsor, Windsor, ON.

- \*Robichaud, L. K. (2004). *Depth of experiencing as a client prognostic variable in emotion-focused therapy for adult survivors of childhood abuse*. (Unpublished master's thesis). University of Windsor, Windsor, ON.
- Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin*, 86, 638-641.
- Rosenzweig, S. (1936). Some implicit common factors in diverse methods of psychotherapy. *American Journal of Orthopsychiatry*, 6, 412-415.
- \*Rudkin, A., Llewelyn, S., Hardy, G., Stiles, W. B., & Barkham, M. (2007). Therapist and client processes affecting assimilation and outcome in brief psychotherapy. *Psychotherapy research*, 17(5), 613-621.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420-428.
- Silberschatz, G., Fretter, P. B., & Curtis, J. T. (1986). How do interpretations influence the process of psychotherapy? *Journal of Consulting and Clinical Psychology*, 54(5), 646-652.
- Stewart, L. A., & Tierney, J. F. (2002). To IPD or not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. *Evaluation & The Health Professions*, 25(1), 76-97.
- Sutton, A. J., & Higgins, J. P. T. (2008). Recent developments in meta-analysis. *Statistics in Medicine*, 27, 625-650.
- Thompson, S. G., & Higgins, J. P. T. (2002). How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine*, 21, 1559-1573.



- \*Toukmanian, S. G., Jadaa, D., & Armstrong, M. S. (2010). Change process in clients' self-perceptions in experiential psychotherapy. *Person-Centered and Experiential Psychotherapies*, 9(1), 37-51.
- \*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.
- Watson, J. C., & Greenberg, L. S. (1996). Pathways to change in the psychotherapy of depression: Relating process to session change and outcome. *Psychotherapy*, 33(2), 262-274.
- \*Watson, J. C., McMullen, E. J., Prosser, M. C., & Bedard, D. L. (2011). An examination of the relationships among clients' affect regulation, in-session emotional processing, the working alliance, and outcome. *Psychotherapy Research*, 21(1), 86-96.
- Whelton, W. J. (2004). Emotional processes in psychotherapy: Evidence across therapeutic modalities. *Clinical Psychology and Psychotherapy*, 11, 58-71.
- Wood, J. (2008). Methodology for dealing with duplicate study effects in a meta-analysis. *Organizational Research Methods*, 11(1), 79-05.
- Zakzanis, K. K. (2001). Statistics to tell the truth, the whole truth, and nothing but the truth: formulae, illustrative numerical examples, and heuristic interpretation of effect size analyses for neuropsychological researchers. *Archives of Clinical Neuropsychology*, 16, 653-667.

## APPENDIX A

### List of contacted researchers

Beck, A. Chicago Group Development Research Team.

Burgess Moser, M. University of Ottawa.

Caspar, F. University of Bern.

Ellis, L. Chicago School of Professional Psychology.

Grindler-Katonah, D. Argosy University American School of Professional Psychology.

Hakim, L. Z. University of Windsor.

Johnson, S. University of Ottawa.

Klug, G. University of Zurich.

Kray, T. North Central Bronx Hospital.

Levitt, H. University of Massachusetts.

Llewelyn, S. University of Oxford.

Pachankis, J. Yeshiva University.

Paivio, S. University of Windsor.

Pos, A. York University.

Ralston, M. University of Windsor.

Robichaud, L. University of Windsor.

Rudkin, A. University of Southampton.

Stiles, W. University of Ohio.

Toukmanian, S. York University.

Vanaerschot, G. University of Antwerp.

Wongpakaran, T. Chiang Mai University.

## APPENDIX B

Studies presented by data taken from research centers

York University	OISE	University of Ottawa	NIMH**
Pos & Escobar (2011)	*Watson et al. (2011)	*Makinen & Johnson (2006)	Missirlian (2011)
*Pos et al. (2009)	Watson & Bedard (2006)	Makinen (2005)	*Hakim (2010)
Goldman et al. (2005)			Jackson (2011)
Pos et al. (2003)			
Levitt et al. (2000)			
Watson & Greenberg (1996)			

---

\*Chosen as primary study for this meta-analysis.

\*\*Studies that used NIMH data set also used York University data set, but only NIMH data was used for the meta-analysis.

## VITA AUCTORIS

NAME: Nikita Yeryomenko  
PLACE OF BIRTH: Frunze, USSR  
YEAR OF BIRTH: 1986  
EDUCATION: York University, B.A., Toronto, ON, 2006-2009  
University of Windsor, M.A., Windsor, ON, 2010-2012