DET PSYKOLOGISKE FAKULTET

Ψ

The association between outcome expectancies and clinical outcomes in psychotherapy
A follow-up systematic literature review

HOVEDOPPGAVE
profesjonsstudiet i psykologi

Henrik Lunde Sagen

Vår 2024

Foreword

One of the more profound realizations I have had during my past five and a half years enrolled at the Professional studies in psychology at the University of Bergen is that the role of a psychologist arguably is less about helping patients and has more to do with helping patients help themselves. While this juxtaposition might seem a bit fraught, its implications are significant from my point of view. This realization helped pique my interest into what can broadly be described as expectation effects, that is how one's positive or negative preconceptions regarding the efficacy of any given treatment or a therapist might potentially affect how the therapeutic processes transpires or to what degree its outcome can be said to be successful. As luck would have it a past colleague of mine, Anders Bolstad, wrote a systematic literature review on the topic as his main thesis spanning the years 2010 to 2017. With the scientist practitioner model in mind merging my clinical experiences and interest with the re-application of a tested methodology for the period following the aforementioned systematic literature review seemed like a prudent approach given the scope and aim of the present paper.

Sogndal 15.05.2024

Henrik Lunde Sagen

Supervisor: Endre Visted

Abstract

This follow-up systematic literature review seeks to examine if research published from March 2017 to present indicates that there is significant relationship between outcome expectation and outcome (Constantino et al., 2011), and to explore if the focus in said research has shifted by comparing the results in the present and Bolstad's (2017) preceding systematic literature review. Eight articles met the selection criteria, where five fully and two in-part supported Constantino et al.'s results. The research on said association can arguably be said to be maturing where one can discern a shift from studies primarily examining *if* there is an association between outcome expectancies and the clinical outcomes to currently exploring the mechanisms active in or acting on such a link. Such a shift arguably re-emphasizes the significance of standardized procedures for how and when outcome expectancies are screened for and kept track of during treatments, to ensure that different studies are screening for the same construct. Having a greater understanding on sample compositions, increasing sample sizes, conducting more treatment-specific qualitative research prior to follow-up larger-sample studies, or facilitating more research in countries with universal health care could be needed to enable future research on potentially more dynamic and nuanced mechanisms.

Abstrakt

Denne systematiske litteraturgjennomgangen søker å utforske om forskning publisert fra mars 2017 til i dag indikerer at det er en signifikant sammenheng mellom utfallsforventinger og kliniske utfall (Constantino et al., 2011), samt å utforske om fokuset i forskning på området har endret seg ved å sammenligne resultatene fra denne med resultatene i Bolstads (2017) forutgående systematiske litteraturgjennomgang. I alt åtte artikler møtte seleksjonsskriteriene, der alle fem av studiene helt og to av studiene delvis støttet Constantino et al. sine funn. Ved at man kan påvise et mulig overordnet skifte fra studier som primært undersøker om det er en sammenheng mellom utfallsforventninger og kliniske utfall til forskning som utforsker mekanismene som er aktive i eller virker på en slik sammenheng kan det fremstå som forskningsfeltet er under modning. En endring til å fokusere på mer dynamiske og nyanserte mekanismer øker sannsynligvis viktigheten av å standardisere prosedyrer for hvordan og når utfallsforventninger blir målt før og under behandlinger, særlig for å sikre at ulike studier kartlegger det samme fenomenet. Det kan og øke viktigheten av å ha en mer nyansert forståelse for utvalgene som benyttes i fremtidige studier.

Contents

Background

Expectations and stigma in psychotherapy

There are some indications that seeking help in the case of mental illness is more associated with stigma than is the case with somatic illness (Mannarini & Rossi, 2019). Research also indicates that there might be a shift in attitudes across generations toward seeking help with some mental illnesses, with it being more likely that younger individuals will seek help (Pescosolido et al., 2021). Oexle et al. (2018) demonstrated that internalization of negative stereotypes defined as self-stigma could hinder recovery by undermining the patient's belief in their self-efficacy. As such, treating mental illness can be said to be associated with more negative expectations than somatic illnesses with regards to not only help-seeking behavior being lessened, but how potential patients evaluate their own ability to benefit from it.

The importance of understanding the role of outcome expectations in therapy

The forementioned potential barriers to seeking or optimally benefitting from therapeutic treatment in the case of mental illness outside of the factors pertaining to the therapeutic methods employed themselves have been the focus of much research. A main yield from such research is the emergence of what is referred to the common factors of psychotherapy, that is – factors such as the real relationship, expectations, and specific ingredients (Wampold, 2015) that are found to be present across all types of psychotherapy. The common factors of therapeutic alliance, therapist empathy, positive regard, genuineness, and client expectations have not only been researched the most, but has also been demonstrated to improve treatment outcomes when they are present to a strong degree (Browne et al., 2021) Research on how these factors can be said to influence therapy can reasonably be posited to even more important in the case of mental illness compared to somatic illness due to the aforementioned potential barriers.

Focusing on the role of expectations specifically, when reviewing the research on the topic it quickly becomes apparent that the research on the topic is a bit fraught. For instance, expectations are operationalized across different studies in quite a few different terms, such as role expectancy, treatment credibility, self-efficacy expectancy, treatment expectancy, outcome expectancy or duration expectancy. While these terms also indicates that the research is seeking to explore how different aspects of expectations can be said to occur in or influence therapy, the lack of a uniform conceptualization also arguably makes it challenging to merge these disparate conceptualizations into a coherent whole. Also, as Wampold (2015) explains, "creating the expectations, through explanation of the patient's disorder, presenting the rationale for the treatment, and participating in the therapeutic actions, is part of therapy." This thus raises the additional issue of when outcome expectations are to be measured – before the onset of the treatment or after the initial session(s). Another hurdle is how expectation is to be measured as there is currently no universally adopted questionnaire.

Bolstad (2017) outlined that the research on how outcome expectancies and treatment outcomes are linked differs across studies in his systematic literature review on the topic. For instance, he presented that while Goldstein & Shipman (1961) and Sotsky et al. (1991) found that higher outcome expectancies predicted more beneficial outcomes in therapy, that Greer (1980) found in a limited study (n=47) that lower expectations yielded more favorable outcomes. Noble, Douglas and Newman (2001) described the relationship to curvilinear – that is lower and higher expectations yielded less favorable outcomes than those who were in-between. The findings are thus not consistent across the studies. Constantino et al. (2011) demonstrated a significant, but weak correlation between outcome expectations and symptomatic recovery in psychotherapy in meta-study based on 46 independent samples (N=8016). Bolstad found that of the nine studies published after Constantino et al. meta-analysis to March 2017 that:

"...five studies that found a significant positive relationship between outcome expectancy and outcome, where two of those only found partial support. Among the remaining studies, four found no relationship and one found a negative relationship (low expectations, better outcomes). (Bolstad, 2017)

Objectives for the present paper

The present paper seeks to continue where Bolstad's (2017) systematic literature review left of. As such, the present paper is also not only obliged to adopt his methodological approach, but also to adopt his framing of the matter to be explored. Bolstad's starting point for his thesis were four logical deductions presented in an article written by Irving Kirsch in 1997 – as presented by Bolstad in the following manner;

1) *The placebo effect occurs in psychotherapy and might be contributing to a significant degree when psychotherapy yields positive outcomes.*

2) *We will have expectations of how we will respond to a stimulus in those instances where we have either information concerning or direct experience with a given stimulus. These expectations are thus linked to our automated reactions which are outside of our ability to consciously control.*

3) *As such, Placebo effects are due to response expectations that takes the form of self-fulfilling prophecy.*

4) *The preceding three claims yields that expectations concerning the outcome of psychotherapeutic treatment are of great importance for the outcome of the treatment. It should thus be feasible to demonstrate that there is a significant positive relationship between these two variables: outcome expectation and outcome.* (Bolstad, 2017)

Similarly to Bolstad's systematic literature review, exploring the available research demonstrating the fourth premise is the aim of this thesis.

Bolstad's (2017) aim was to explore how any research published after Constantino et al. published their comprehensive meta-analysis "Expectations" on the subject in 2011 shed light on the matter. Like Constantino, Bolstad found that the sum of the research indicated the link to be tentative at best. Constantino et al. postulated that a reason behind the weak link between outcome expectancy and the outcome might have been that publication bias prevented studies that indicated that there was a lack of connecting – be it statistically significant or not - between the two factors preventing the study from being published. Such considerations cannot easily be addressed as one do not typically gain access to unpublished studies when conducting meta-analysis on a given subject. Publication bias could also be in effect if the studies in question lack the methodological stringency required to be published, be it due to too small a sample, the way the data collection has been organized or if the operationalization of the thesis or definition of the study's core constructs were done in a manner not in keeping with how the potential peer-reviewers assessed as satisfactory.

Bolstad also raised some other concerns, namely the lack of a uniform way of measuring the outcome expectancy at the start of the studies, both with regards to when it is measured or how. The studies in Bolstad's systematic literature review did not all measure their outcome expectancy associated with their participants before the onset of the study. Several of the studies took this measurement after the initial patient-therapist interaction, raising the concern that interaction with their respective therapists can skew the measurements. How the measurement was conducted also differed greatly between the studies, ranging from asking a couple of study-specific question to administering a vetted questionnaire. As such, no standardized way was used

across the studies. This could also mean that the outcome expectancies in themselves were so unrealistic that what one in the end ended up demonstrating was not how the two variables were linked, but that the outcome expectancy in and of itself was outside of anything one could realistically hope to achieve.

Furthermore, how two of the studies in Bolstad's systematic literature review computed the data statistically could also have contributed to the less than clear saliency of the link between outcome expectation and the outcome itself. By using several independent variables that interact in iterative steps within the statistical analysis the relative strength of one or more of these might be skewed.

As such, the present study aims to continue were Bolstad's work ended, with the considerations outlined above in mind. There are several methodological and practical reasons for this choice. First and foremost, the approach is proven to yield results. Secondly, the repetition of the same approach for a new period could enable a direct comparison on the research within the field between different periods in time. This can for instance potentially enable the determination if the research has shifted focus over time, if it has become more nuanced in what it is that it seeks to explore or if there are potential short comings that mainly become apparent when comparing research over time. Examples of the latter can be inconsistent use of metrics, measurement procedures or definitions. Thirdly, as the present paper like Bolstad's (2017) is a thesis for postgraduate course in the Professional Studies in Psychology at the University in Bergen, the approach is also proven to be commensurate with the scope of the present task.

Methods

The aim of this systemic review is to use the same methodology as used in "Sammenhengen mellom utfallsforventninger og utfall av psykoterapi: En systematisk litteraturgjennomgang" (Bolstad, 2017) the period following Bolstad's work. As such this systematic literature review focused on studies published from March 1st, 2017, until the 11th (PubMed) and 12th (PsycINFO) of April 2024 respectively. The period in question is thus from roughly half a month after the period addressed in Bolstad's systematic literature review to the present. This follow-up systematic literature review also uses the same definition of outcome expectations that Bolstad used in his systematic literature review, as defined by Constantino et al. as something that "… reflect patients' prognostic beliefs about the consequences of engaging in treatment…" (2011, s. 3).

Several steps were enacted to minimize the influence of selection bias in the selection process. The most important of these was the adherence to the best practices for conducting a systematic literature review as defined by referred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) (Page et al., 2020), and secondly to execute a clearly defined strategy with regards to the search as well as having pre-defined selection criteria. However, the methodological approach as defined by Bolstad's previous systematic review do entail some significant short comings that are outlined later in this paper.

Selection of databases and studies

PsycINFO and PubMed were utilized for the searches. The aim of the search was thus to highlight research on the matter at hand published in the period after Constantino et al. (2011)

and Bolstad (2017) as made possible by Bolstad's methodology, not alter Bolstad's methodology. This was done to ensure a more holistic approach.

The selection criteria for this systematic review are largely the same as in Bolstad's systematic literature review in that the studies had to

> *...investigate the connection between outcome expectations and symptomatic recovery, b) use measures of the patient's own outcome expectations before starting treatment or after the first meeting with the therapist, c) use measurements of symptomatic recovery immediately after treatment, d) involve treatment in form of psychotherapy, ... , f) use samples consisting mainly of adult patients...* (Bolstad 2017)

However, a couple of Bolstad's criteria were adopted to better accommodate the present study, namely his criterium "e) be published after 1 January 2010" (Bolstad 2017) was altered to reflect the current period starting from the 1st of March 2017 while his criterium "g) be published in English, Norwegian, Swedish or Danish" (Bolstad 2017) was moderated to only include studies published in English due to the way the search was conducted. His remaining selection criteria outlined below were adopted with only a few exceptions as specified below.

> *Studies were excluded if studies had used: a) samples consisting of non-clinical groups (e.g. students), b) non-clinical outcome measurements (e.g. physical activity), c) non-clinical forms of treatment (e.g. light therapy and self-help) or d) measures of outcome expectations combined with measurements of adjacent phenomena (e.g. "treatment expectations", "treatment credibility" or "role expectations"). Inclusion and exclusion criteria are inspired by, but not identical to, those used by Constantino et al. (2011). Studies were not excluded based on its country of origin.* (Bolstad 2017)

Furthermore, a few criteria were added. Systematic literature reviews or meta studies published on the subject in the period in question were excluded. The choice was also made to stringently adhere to the results as yielded through the database searches, not to explore other search strategies such as including papers citing the Constantino et al. (2011) meta-analysis. Studies outlining expectations to a given part of the treatment and not to its outcome were also excluded. A common example of such a paper would be research on expectation violation as driver for the beneficial outcomes in exposure therapy, studies focused on expectations with regards to what was going to transpire in the exposure treatment specifically and not on expectations as defined to. be linked to the overall outcome of the therapy in itself, or studies that found that manipulating the patient's belief in the efficacy of the expectation effect underway in the treatment was what predicted the correlation or perceived causality (or lack there-of) of the pre- and post-measurements. Studies whose patients were either children or adolescents were also excluded. The reason for this exclusion was that their parents were often included as a dyad (a single parent and the child) or a triad (both parents and the child) instead of focusing on the expectation of the child or adolescent as a standalone patient. The reason that studies focusing on adolescents as standalone individuals also were excluded is because most treatments focus on the adolescent as a part of a family system where the family system is the "true" focus of the treatment, and less so the adolescent as a standalone individual as is seen adult treatments. Studies focusing on the expectation of the therapist rather than those of the patients were also excluded as the aim of the present paper is not to explore second order effects or indirect causes that might affect the patient, but how processes related to the patient's outcome expectancy is associated with the outcome. Studies involving treatments like electroconvulsive therapy were also excluded as treatment is administered while the patient is sedated.

Keywords and phrases

The keywords, phrases and search fields were adopted from Bolstad's systemic literature review with some minor alterations. These are presented in Table 1. The search in PsycINFO was amended in that Bolstad's phrase "therapy outcome" was altered to "therap* outcome". This was done to broaden its potential yield. The search in PubMed was altered to excluded "subject heading word" as a search field, as neither search term used in the PubMed search were defined as a subject heading word and as such would not have altered the yield of the search.

Table 1
Database and search criteria

Database and period	Platform	Search fields	Keyword or phrase
PsycINFO 2017-2024	OvidSP	Title, abstract and "Subject Headings"	"expecta*" AND "treatment outcome" OR "therap* outcome"
PubMed 2017-2024	OvidSP	Title and abstract	"expecta*" AND "therapy outcome"

Collection of data

The collection of data was conducted during two searches, one in the database PsycINFO and the second in the database PubMed. The screening and selection of relevant studies process consisted of two steps. The first step was to read the tittles and abstracts of all the articles yielded from the two searches and screen them for articles that were within the selection criteria of the present study. The second step was to read the full-text articles identified in the first step and select the articles that met the selection criteria of the present study.

Results

The initial yield from the searches tallied 182 articles, consisting of 157 from PsycINFO and 25 from PubMed. The search on PubMed was conducted on the 11th of April, while the PsycINFO search was conducted on the 12th of April.

The 182 articles were screened for relevancy based on their titles and abstracts. After reviewing their titles and abstracts 124 articles, 14 papers, three meta studies and 12 reviews were excluded based on the selection criteria. The 124 articles included 10 studies that were duplicates found in both searches, two were republications of the same study at different times, and two were corrective republications due to errors having been discovered in figure in both articles necessitating a re-publication. The initial screening process was repeated in full two weeks after the initial screening to prevent selection bias or inconsistent selection from interfering in the selection process. The difference between the yield after the two initial screening processes were a total of six articles – two "extra" in the first and four in the second, in addition to 22 that were selected in both iterations of the initial screening. Please not that all of these six were excluded in the following step. As such, the end effect of repeating the initial screening process was minimal.
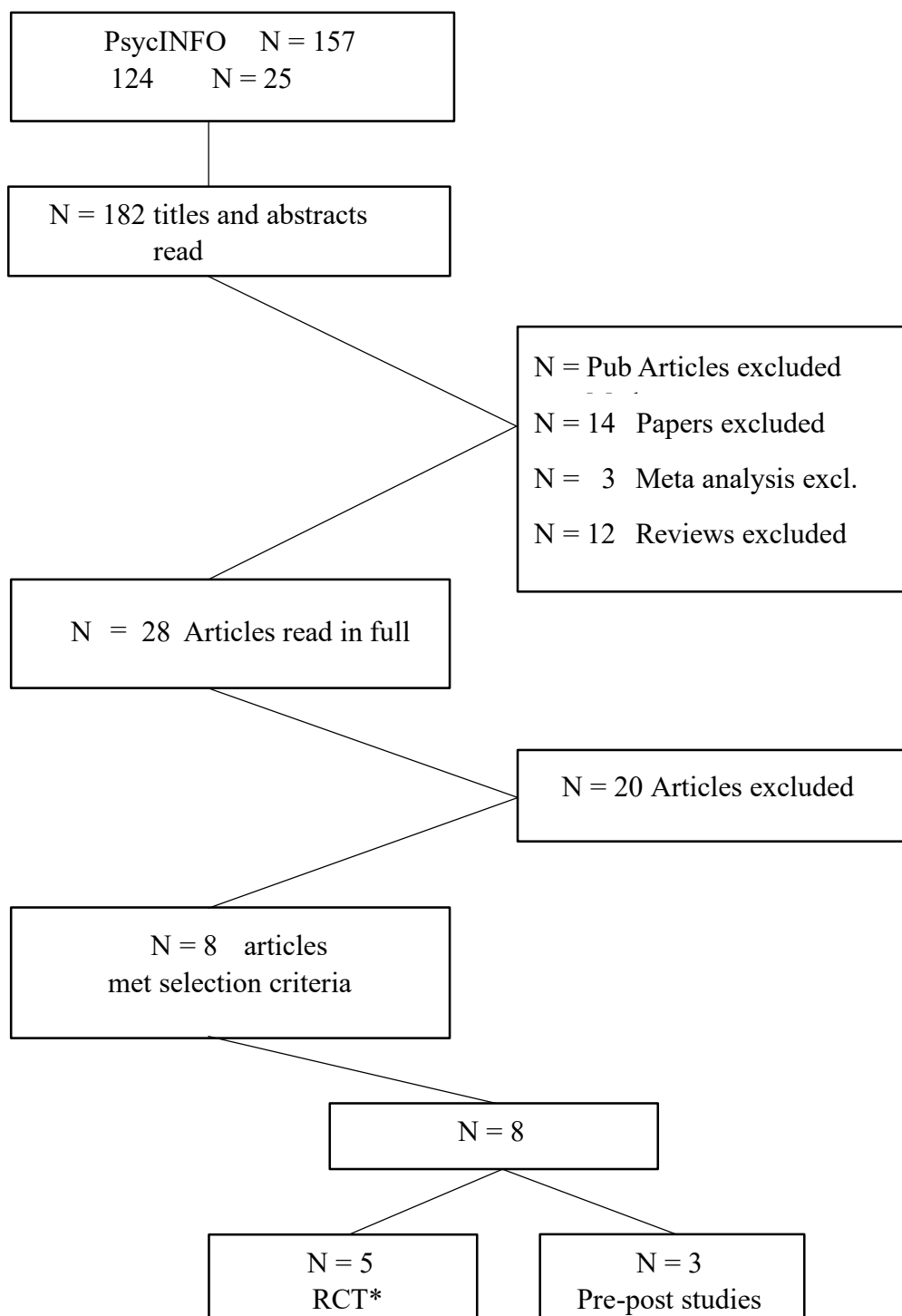
The next step included doing a more comprehensive screening by reading the full text article. The initial method was to screen the method sections for indications that the article did not comply with the defined selection criteria of the present study. Examples of such exclusion criteria were that the participants were of a type outside of the remit of this paper such as adolescents, children or parent-child constellations, that the instrument used to measure the expectancy outcome was not a questionnaire of proven reliability or more than a single question that only loosely could be said to screen for outcome expectancy as defined by Constantino's

(2011) definition outlined previously, or that the aim and results of the study were not relevant for the objectives in this follow-up systematic literature review. Other reasons for exclusion were that how the study had defined outcome expectancy had more to do with either factors occurring within the treatment itself such as expectancy violation in exposure therapy, or that the outcome measurement mainly was defined by outcome satisfaction defined as a highly subjective, non-generalizable metric specific to the patient in question measured by one or several questions specific to that study and not a documented questionnaire. This step led to the exclusion of an additional 20 articles.

Eight articles met the selections criteria outlined above. These are presented in Table 2. The screening process is presented in Figure 1.

Figure 1.

Identification and selection of relevant articles

| PsycINFO   N = 157 |
| 124   N = 25 |

N = 182 titles and abstracts
read

N = Pub Articles excluded

N = 14   Papers excluded

N =   3   Meta analysis excl.

N = 12   Reviews excluded

N = 28 Articles read in full

N = 20 Articles excluded

N = 8   articles
met selection criteria

N = 8

| N = 5 | N = 3 |
| RCT* | Pre-post studies |

Note. *Randomized controlled trials

Three of the studies were conducted outside of the US, one in Canada (Visla et al., 2016) one in Germany (Terock et al., 2018) and one in the United Kingdom (Turner et al., 2019). Five of the studies used a randomized control trial (RCT) design, while the remaining three used a pre-post type study design. The number of participants in the studies were quite uniform across the various studies ranging from 91 in the Canadian study (Visla et al., 2016) to 175 in one of the studies from the US (Beck et al., 2021), with the German study (Terock et al., 2018) being the outlier at 518 participants. Except for Beck et al. (2021) and Gobin et al., (2018) whose study were all male and female respectively, and Visla et al. (2016) and Sauer-Zavala et al. (2018) which did not specify the composition of their participants with regard to age or sex, the remaining studies all had female inclusion ranging from 43% (Strauss et al., 2018) to 95% (Turner et al., 2019). Of these five studies, four consisted of more than 65% female patients, which can be said to be a bit skewed. Similarly, the mean age was relatively uniform across the various studies at roughly 40 years, with Beck et al. (2021) and Turner et al. (2019) studies' mean age of 54.9 years and 28.8 years respectively, and the two aforementioned studies that did not specify the age composition of its participants being the outliers. All the studies addressed a single disorder as the main disorder. All but two of the studies catered to at least two different types of treatment.

All but one study measured the outcome expectancy using different standardized questionnaires. The exception to this was Strauss et al. (2018) that measured the patients' expectancies regarding improvement in treatment on obsessions, compulsions and general distress through a study specific question asked after the second session using three Likert scales (0–8) specific to each dimension where a higher score indicated a higher expectancy. that were summarized to form the patients' expectancy ratings (PER). Chambless et al. (2017) and Sauer-Zavala et al. (2018) both used the Credibility/Expectancy Questionnaire (CEQ) which has a total

of six question, three measuring cognitively based credibility of the treatment in question with the remaining three measure the relatively more affectively based expectancy (Devilly & Borkovec, 2000). Beck et al. (2021) used the Expectancy Rating Questionnaire (ERQ which assesses if the treatment seemed logical at treatment onset and the degree of confidence that the participant had in the intervention through four questions using a 0-9 scale where the final score is the sum of these and a higher score is indicative of a higher expectation (Borkovec & Nau, 1972). Gobin et al. (2018) used the four item Treatment Expectancy Questionnaire (TEQ) at Session 2 to screen the patients' perception on treatment credibility and expected outcome (Borkovec & Nau, 1972). Terock et al. (2017) used Questionnaire for the Measurement of Psychotherapy Motivation (FMP-30) that through 30 question measures the degree of suffering, psychotherapeutic treatment expectations, psychosocial lay etiology, and secondary gains of illness at treatment onset (Kraft et al., 2015). Turner et al. (2019) used the Expectation of Improvement and Suitability of Treatment Form (EIST) at session six to assess the patient expectations and suitability of treatment, each using a 15-point scale (0 -15) where higher scores indicate higher assessed expectations and suitability (Agras et al., 2000). Visla et al. (2018) used the Outcome Expectancy Scale (OES) to assess patients' expectation for improvement using a 5-point Likert-type scale (1 - 5) on three questions indicating more positive outcome expectations through higher scores (Ogrodniczuk & Sochting, 2010).

The studies all measured the outcome of the treatment through different questionnaires. The variation in method is likely primarily due to the differences in diagnosis being treated. All used standardized questionnaires that are commonly and routinely used in clinical settings to variably assess symptom severity both at the onset, during and after the treatment, and are as such not presented in detail.

Table 2

Description of the included studies

Author / year / country	Design	Participants*	Metric OE	Disorder	Treatment	Metric outcome
Beck et al. / 2021 / USA	RCT	175/54.9/0 %	ERQ	PTSD	GCBT / GPCT	PCL-5
Chambless et al. / 2017 / USA	RCT	161/39.4/65.0 %	CEQ	Panic Disorder	PFPP / CBT	PDSS
Gobin et al. / 2018 / USA	RCT	126/42.1v 47.3c/ 100 %	TEQ	PTSD	CPT	CAPS, PTCI
Sauer-Zavala et al. / 2018 / USA	Pre-post study	179/ns/ns	CEQ	Anxiety	UP, SDP	HARS, WAI-S
Strauss et al. / 2018 / USA	RCT	98/39.2/43 %	PER	OCD	EX/RP & SMT	QLESQ, YBOCS, WAI-C
Terock et al. / 2017 / Germany	Pre-post study	518/43.1/70.2 %	FMP-30	Alexithymia	CBT / PDT	TAS-20, SCL-90-R
Turner et al. / 2019 / UK	Pre-post study	128/28.8/95 %	EIST	Eating Disorder	CAT / CBT / PDT	EDE-Q, PBQ-SF
Visla et al. / 2018 / Canada	Pre-post study	91/ns/ns %	OES	Depression	CBT	BAI, BDI-II, IIP-28, WAI

*Note. *Participants (N)/mean age / percentage women. BAI - Beck Anxiety Inventory; BDI-II – Beck's Depression Inventory-II; CAT – Cognitive Analytic Therapy; CAPS- Clinician-Administered PTSD Scale; CBT – Cognitive Behavior Therapy; CEQ - Credibility/Expectancy Questionnaire; CPT - Cognitive Processing Therapy; EDE-Q - The Eating Disorders Examination–Questionnaire; EIST – Expectation of Improvement and Suitability of Treatment Form; ERQ- The Expectancy Rating Questionnaire; EX/RP - Exposure and response prevention; FMP-30 -Questionnaire for the Measurement of Psychotherapy Motivation; GCBT - Group Cognitive-Behavioral Treatment; GPCT - Group Present Centered Therapy; HARS - Hamilton Anxiety Ratings Scale; IIP-28 - Inventory of Interpersonal Problems – 28i; ns – not specified; OE- Outcome Expectancy; OES – Outcome Expectancy scale; PER – Patients Expectancy Ratings; PBQ-SF - The Personality Belief Questionnaire – Short Form; PCL-5 - The PTSD Checklist for DSM-5; PDSS – Panic Disorder Severity Scale Self-report; PDT – Psychodynamic Therapy; PTSD – Post Traumatic Stress Disorder; QLESQ - Quality of Life Enjoyment and Satisfaction Questionnaire; SMT - Stress management training; SDP- Single Diagnosis Protocols; SLC-90-R - Symptom Checklist-90-Revised; TAS-20 - Toronto Alexithymia Scale 20; TEQ - Treatment Expectancy Questionnaire; WAI – Working Inventory Scale; WAI-C - Working Inventory Scale-Clients; WAI-S- Working Alliance Inventory-Short Form; YBOCS - Yale-Brown Obsessive Compulsive Scale for OCD.*

Randomized Controlled Studies

Five of the articles that met the inclusion criteria used a randomized control trial (RCT) study design. A RCT study design randomly assigns participants to two or more separate groups that both undergo an intervention specific to that group (Aggarwal, R., & Ranganathan, P., 2019). The randomization process is meant to ensure that the two groups are comparable such that the outcome is determined by the intervention itself, and not due to characteristics of the participants. The RCT study design is usually considered the most stringent way to establish whether a cause-effect relation exists between the intervention and the outcome (Sibbald & Roland, 1998). The five RCT studies are described below. The results are limited to those relevant to the present study in the written description yet are more expansive in their summarization that is found in Table 3.

Beck et al. (2021) explored five variables[1] for their association with trajectories of change during Group Cognitive-Behavioral Treatment[2] (GCBT) and Group Present Centered Therapy (GPCT). The association with trajectories of change was explored across the five Post-Traumatic Stress Disorder (PTSD) symptom clusters[3] as defined by The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) (American Psychiatric Association, 2013). The study had 175 participants with a mean age of 54.9 that were all men. The group composition for each treatment were comparable. Beck et al. used Unconditioned Growth models and Intercepts-and-Slopes-as-Outcomes to analyze their data. Results indicated potential support for a moderating effect of outcome expectancies on expected change in cognitions/mood (p = .040) GCBT. Greater reductions in cognitions/mood controlling for other patient-level characteristics (b = -.496,

---

[1] intrusion, avoidance, cognitions/mood, and arousal/reactivity
[2] trauma-focused
[3] Pre-Treatment-Post-Traumatic Stress Disorder, pre-treatment depression, age, whether combat was the index trauma, and outcome expectancy

$p < .001$; $d = 1.13$) may be associated with higher expectancies of positive outcome following the first session in GCBT relative to individuals reporting lower expectancies of positive outcome (b = -.219, $p = .120$; $d = .50$). Beck et al. speculated that these discoveries were more noticeable in a trauma-focused treatment such GCBT over GPCT, since trauma-focused GCBT emphasizes learning new skills as well as requiring between-session practice.

Gobin et al. (2018) sought to identify factors that may explain why female veterans showed poorer treatment response than civilian women in a previous RCT study of cognitive processing therapy (CPT). The all-female study had 126 participants, 21 veterans and 105 civilians, with a mean age of 42.1 and 47.3 respectively. The two groups were somewhat similar in terms of demographic characteristics and comorbid depression, if not in size. They used structural equation modeling to investigate the role of seven clinical and treatment variables[4] to explain the reduced treatment response to CPT in veterans compared to civilians. Gobin et al. found that the veterans' smaller reduction in negative posttraumatic cognitions that was found to fully mediate the decreased treatment benefit (lower PTSD symptoms), also reported lower treatment expectancy at baseline than the civilians. Gobin et al. inferred that the lower treatment expectancies in veterans might be related to a lower level of cognitive flexibility, especially as the veterans and civilians did not differ in baseline cognitions.

Sauer-Zavala et al. (2018) primarily sought to determine if the association between patient outcome expectancy and post-treatment outcome is mediated by patient-rated working alliance quality, and if the indirect effect of expectancy on treatment outcome through the quality of the working alliance is moderated by treatment condition. They also sought to explore

---

[4] Variable / measurement: PTSD Symptoms / CAPS; History of interpersonal trauma / 12 items from Life Stressors Checklist—Revised; History of noninterpersonal traumatic and stressful events / 19 items from Life Stressors Checklist—Revised; Depressive symptoms / Beck Depression Inventory-II; Concurrent mental health (MH) services / interview; Treatment expectancy / Treatment Expectancy Questionnaire; Posttraumatic cognitions / Posttraumatic Cognitions Inventory

whether a transdiagnostic CBT protocol (UP) or single diagnosis CBT protocol (SDP) moderated this indirect relationship. The study had 179 participants yet did not provide any additional information on the study's participants such as their age or gender. Sauer-Zavala et al. used Pearson correlations to capture relationships between study variables[5], while the evaluation of the mediation and moderated mediation models was conducted using the PROCESS macro developed by Hayes (2013). The association between expectancies and outcome was found to be partially mediated by the working alliance (B = .037, SE = .05, 95% CI [.005, .096), if only present for SDP patients. The association between expectancies and outcome was not found in the UP patients. As such, Sauer-Zavala specified that while outcome expectancies and working alliance quality may interact to influence treatment outcomes, differing cognitive-behavioral treatment approach might affect the nature and strength of the relationships among these constructs.

Strauss et al. (2018) sought to examine the relative contribution of expectancy and alliance in predicting outcome over and above specific ingredients of treatment for obsessive compulsive disorder (OCD) in two types of CBT treatments: Exposure and Response Prevention (EX/RP) or Stress Management Training (SMT). The study had 98 participants with a mean age of 39.2 years of which 43% were women. Strauss et al. analyzed the data using the full intent-to-treat longitudinal mixed effects models (LMLM) (Chakreaborty, H. & Gu, H, 2009). Although both groups exhibited moderate expectancies, both patients' and therapists' expectancies were significantly higher in EX/RP. The mean patient expectancy in EX/RP was 18.27 (SD = 3.90) compared to 15.57 (SD = 4.64) in SMT ($t97 = 3.16$, p = .002). Strauss et al. examined the time by expectancy interactions to estimate to what degree early expectancies moderate changes in Yale-Brown Obsessive Compulsive Scale for OCD (YBOCS) and Quality of Life Enjoyment and

---

[5] outcome expectancy, working alliance, and treatment outcomes

Satisfaction Questionnaire (QLESQ). No significant association was identified. As such, Strauss

et al. found that expectancies were not related to outcomes, something that they highlighted to be

partially consistent with the literature on OCD (Başoğlu et al., 1988; Freeston et al., 1997;

Steketee et al., 2011) if not with Constantino et al. (2011).

Chambless et al. (2017) explored predictors and moderators of response to psychotherapy

for panic disorder patients with or without agoraphobia in Cognitive Behavioral Therapy (CBT)

and Panic-Focused Psychodynamic Psychotherapy (PFPP). The study had 161 participants with a

mean age of 39.4. 65% of the participants were women. 127 of the participants had a primary

diagnosis of panic disorder, and 21% had agoraphobia as a comorbid diagnosis. The group

composition for each treatment were comparable. Chambless et al. analyzed the change in

treatment over time using a shared parameters approach (Ten Have et al., 1998) instead of

standard approaches to the longitudinal analysis of treatment data due to their study's lack of data

missing at random. Their results indicated that a higher expectancy for treatment gains was

predictive of greater change ($t(84) = -4.85$, $p < .0001$, $d = -1.05$, CI 95% [-1.50, -0.60]). Higher

expectancy also worked as a significant moderators: Patients with low expectancy of

improvement improved significantly less in PFPP than in CBT ($t(83) = 2.39$, $p = .02$, $d = 0.52$, CI

95% [0.09, 0.95]), whereas this was not the case for patients with average or high levels of

expectancy ($ds \leq 0.28$, $ps \geq .20$). Chambless et al. specified that this is consistent with previous

studies (Constantino et al., 2010; Porter & Chambless, 2015). Of note is that the expectancy

measure might have been affected by the patient having met their therapist for one or two session

and being explained the treatment's rationale before being screened.

Table 3
Results for RCT studies

Study	Objectives	Results
Beck et al., 2021	To explore the role of moderators of change in GTO for PTSD through five variables for their association with trajectories of change during GCBT and GPCT (PT-PTSD, PTD, age, whether combat was the index trauma, and OE).	Positive outcome expectancies enhanced reductions in cognitions/mood in GCBT while pre-treatment PTSD moderated avoidance reduction. Change in intrusion and arousal/reactivity symptoms was moderated by PTD in GPCT.
Chambless et al., 2017	To explore predictors and moderators of response to psychotherapy in CBT and PFPP for panic disorder patients with or without agoraphobia.	Higher expectancy for TG and later AoO were predictive of greater change. Both variables were also significant moderators: Patients with low expectancy of improvement improved significantly less in PFPP than in CBT, whereas this was not the case for patients with average or high levels of expectancy. Patients with AoO at ≥27.5 yrs. old panic disorder they fared as well in PFPP as CBT, while CBT was the more effective treatment for low and mean AoO.
Gobin et al., 2018	To identify factors that may explain why female veterans showed poorer treatment response than civilian women in a previous RCT of CPT.	Differences in CAPS scores at posttreatment were largely mediated by negative posttraumatic cognitions, as measured by the PTCI. While both groups had similar PTCI scores at baseline, civilians had significantly lower PTCI scores at posttreatment, which predicted lower CAPS scores at posttreatment. This mediation appeared to be at least in part explained by veterans' lower TE.
Sauer-Zavala et al., 2018	To explore if: (1) the association between patient OE and PTO is mediated by the quality of the WA, as rated by the patient; and (2) the indirect effect of expectancy on TO through the quality of the WA is moderated by TC.	Sauer-Zavala et al. found that the relationship between expectancies and outcome was partially mediated by the working alliance. Within-condition analyses showed that conditional indirect effect was only present for SDP patients. Working alliance did not account for the association between expectancies and outcome the UP condition.
Strauss et al., 2018	To examine the relative contribution of expectancy and alliance in predicting outcome over and above specific ingredients of treatment for OCD in two types of CBT: EX/RP and SMT.	Treatment type was a substantially stronger predictor of symptom reduction compared to alliance and expectancy. However, neither specific nor common factors predicted improvement in quality of life very well. Only in EX/RP, symptom change was associated with subsequent changes in alliance. Finally, therapist effects were negligible.

Note. AoO – Age of Onset; CAPS - Clinician-Administered PTSD Scale; CBT – cognitive behavior therapy; EX/RP - Exposure and Response Prevention; CPT - cognitive processing therapy; GCBT - Group Cognitive-Behavioral Treatment (trauma-focused); GPCT - Group Present Centered Therapy (non-trauma focused); GTO – Group Treatment Outcome; OE – Outcome Expectancy; SMT – Stress Management Training; TC- Treatment Condition; TE – Treatment expectancies; TG – Treatment Gains; PTCI - Posttraumatic Cognition Inventory; PFPP - Panic-Focused Psychodynamic Psychotherapy; PTD – Pre-treatment Depression; PTO – post-treatment outcome; PT-PTSD – pre-treatment PTSD; PTSD – Post Traumatic Stress Disorder; SDP - Single Diagnosis Protocols; UP – Unified Protocol; WA- Working Alliance.

Pre-post studies

Three of the articles that met the inclusion criteria used what can be described as a pre-post study design. A pre-post study design measures one or several variables before and after an intervention has been conducted on a set of participants (Aggarwal, R., & Ranganathan, P., 2019). The pre-post study design's lack of a direct comparison such as administering a different intervention or control on a second subset of participants and the design's limited ability to eliminate extraneous factors such natural variation or co-occurring events means that studies using this design cannot necessarily reliably attribute the outcome to the intervention being tested. As such, the RCT study design presented previously is generally considered a stronger design. The results are limited to those relevant to the present study in the written description yet are more expansive in their summarization that is found in Table 4.

Turner et al. (2019) sought to determine if severity of eating disorder psychopathology (EDP) at treatment onset would predict eating psychopathology (EP) at treatment end, and if personality features (PF) and patient expectation (PE) would impact (EP) at the treatment end over EP severity at treatment onset. The study had 128 participants, of which 95% were women. The mean age was 28.8. While the participants were divided into three groups that received cognitive behavioral therapy (CBT, n=94), cognitive analytical treatment (CAT, n= 22) and psychodynamic treatment (PDT, n=12), the selection for what group each participant would join was not done as a part of RCT study design, but collaboratively to match the patient with the treatment that seemed the best suited for their needs. Multiple regression was chosen to analyze the impact of eating disorder severity, patient expectations, and personality features on eating pathology at the end of treatment. Turner et al. demonstrated strong patient expectations in relation to suitability and success of the delivered treatments and found that the end of the

treatment outcome scores were associated with the initial expectancy scores ($t = 2.80$, $P = .007$). Turner et al. summarized their results in that:

> *"...patients' expectations regarding the likely success of therapy have a significant impact on treatment outcome. Patients who develop a strong belief that treatment will work may be more likely to actively engage in the process of therapy to achieve their recovery goals."* (Turner et al., 2018)

Visla et al. (2018) sought to explore if early to mid-treatment alliance quality would mediate the relation between baseline outcome expectancy and post-treatment outcome; if mid-treatment alliance would mediate the relation between early-treatment outcome expectation and post-treatment outcome; if early-treatment outcome expectation would mediate the relation between early alliance and post-treatment outcome; and if patients with higher levels of baseline outcome expectancy would form better early with their therapists, which in turn would relate to favorable subsequent outcome expectancy and ultimately post-treatment outcome. The study had 91 participants, whose mean age or gender was not specified. The study used simple and serial multiple mediation analyses using a wide range of added steps to ensure that the analysis be suitable to the specifics of their data and the objectives of the study. These added steps are explained in full in their study. Visla et al. found that the relations between baseline outcome expectancy and both post-treatment anxiety and post-treatment depression were mediated by alliance quality. They also found that the early-treatment outcome expectation – post-treatment anxiety relation was mediated by mid-treatment alliance and that the relation between earlier alliance and post-treatment interpersonal problems was mediated by the during-therapy outcome expectation. The relation between baseline outcome expectancy and post-treatment interpersonal problems was mediated by earlier alliance and during-treatment outcome expectation acting in

turn. No other tested model was significant. As such, Visla et al. found that outcome expectation and alliance have a bidirectional relation, and that both directions influence the outcome.

Terock et al. (2017) sought to determine if and the dimensions of psychotherapeutic treatment motivation predicted the treatment outcome in terms of symptom reduction, if alexithymic personality features were correlated with the dimensions of psychotherapeutic treatment motivation; and if alexithymia and dimensions of psychotherapeutic treatment motivation interact in their effects on treatment outcome. The study had 518 participants of which 70.2% were female. The mean age of the participants was 43.1 years. The study analyzed the data using hierarchical linear regression analysis, partial correlation analyses as well as conditional regression analyses. Terock et al. found that TAS-20 ($\beta = 0.08$; $p = 0.035$) and treatment expectations ($\beta = -0.100$; $p = 0.009$) were significant predictors of global symptom severity at treatment end. Partial correlation analyses of TAS-20 and the FMP subscales, adjusted for age, sex, and global symptom severity at admission, revealed significant negative correlations of TAS-20 and treatment expectations ($r = -0.116$; $p = 0.008$), indicating that perceived suitability and expected outcome of the treatment decrease with increasing alexithymia scores. Terock et al. emphasized the particular importance of enhancing treatment expectations in alexithymic patients.

Table 4
Results for pre-post studies

Study	Objectives	Results
Terock et al., 2017	To explore if: (1) alexithymia and the dimensions of PTM predicted the TO in terms of symptom reduction; (2) alexithymic PF were correlated with the dimensions of PTM; and (3) alexithymia and dimensions of PTM interact in their effects on TO.	Terock et al. found that TAS-20 and TE were found as significant predictors of GSS at TE. Partial correlation analyses of TAS-20 and the FMP-30 subscales, adjusted for age, sex, and GSS at admission, revealed significant negative correlations of TAS-20 and TEX, indicating that PS and EO of the treatment decrease with increasing alexithymia scores.
Turner et al., 2019	To explore if: (1) severity of EDP at treatment onset would predict EP at treatment end; (2) if PF and PE would impact EP at the treatment end over EP severity at treatment onset.	Turner et al. found that: ED severity significantly predicted EP at treatment end. However, strength of PoE regarding treatment success significantly predicted EP over and above baseline level of ED severity. PF at baseline were not significantly related to treatment outcome.
Visla et al., 2018	To explore if: (1) early to mid-treatment AQ would mediate the relation between baseline OE and PTO; (2) MA would mediate the relation between ETOE and PTO; (3) ETOE would mediate the relation between EA and PTO; and (4) patients with higher levels of BOE would form better EA with their therapists, which in turn would relate to favorable subsequent OE and ultimately PTO.	Visla et al. found that: (1) The relations between BOE and both PA and PD were mediated by AQ; (2) the ETOE PA relation was mediated by MA; (3) the relation between EA and PIP was mediated by DTOE; and (4) the relation between BOE and PIP was mediated by two variables acting in turn, EA and DTOE. No other tested model was significant.

Note. AQ – Alliance Quality; BOE – Baseline Outcome Expectancy; DTOE - During-Therapy Outcome Expectation; EA – Earlier Alliance; ED – Eating Disorder; EDP - Eating Disorder Psychopathology; EO – Expected Outcome; EP - Eating Psychopathology; ; ETOE - Early Treatment Outcome Expectation; FMP-30 - Questionnaire for the Measurement of Psychotherapy Motivation; GSS - Global Symptom Severity; MA – Mid-Treatment Alliance; OE – Outcome Expectantcy; PA – Post-treatment Anxiety; PD – Post-treatment Depression; PE – Patient Expectation; PF – Personality Features; PIP – Post-treatment Interpersonal Problems; PoE - Positive Expectations; PTO – Post-Treatment Outcome; PS - perceived suitability; PTM - Psychotherapeutic Treatment Motivation; TAS-20 - Toronto Alexithymia Scale 20; TC – Treatment Condition; TE - Treatment End; TEX – Treatment Expectations; TO – Treatment Outcome.

Results summarized

Seven of the articles demonstrated a significant association between outcome expectancies and the clinical outcome in at least one of the treatments being explored in the study in question. Of these, five articles demonstrated a direct significant association. Beck et al. (2021), Turner et al. (2019), Gobin et al. (2018), Chambless et al. (2017) and Terock et al. (2017). Sauer-Zavala et al. (2018) and Visla et al. (2018) demonstrated it implicitly by demonstrating an outcome expectancy – clinical outcome association by exploring how said association was mediated by factors such as during treatment expectations and working alliance. Sauer-Zavala et al. (2018) and Strauss et al. (2018) found that expectancies were not related to outcomes, in one of the two and in both of two treatments in addressed in their studies respectively.

Discussion

As specified previously, the present paper has two main aims. Firstly, to examine if the current research indicates that there is significant relationship between outcome expectation and outcome. Secondly, to explore what, if any changes have occurred in the focus of the research over time by comparing the research from the period detailed in Bolstad's (2017) systematic literature review to that of the period of the present study. A way of discussing any differences that are found between the periods could be to focus on observable concerns that would potentially need to be addressed going forward, that is concerns that are arguably made apparent when scrutinizing research effort when the two periods are seen as a whole that not necessarily is apparent when scrutinizing any one study. Please note that the second objective will be given the most focus in the following discussion, as the main points Bolstad (2017) made regarding the first objective still stands. Bolstad's careful deliberation arguably adequality addresses the results

from the articles included in the present follow-up systematic literature review in terms of an association between outcome expectancies and clinical outcomes as the nature of the outcome expectancy – clinical outcome has not changed in psychotherapy in the intervening years. Please note that this is despite the fact of Bolstad not finding support for Constantino et al.'s finding (2011) and the present study doing so. The present study thus is neither focused on bolstering what Bolstad presented by restating similar observations with examples from the articles in the present study, nor adding new contextual representations of the similar arguments by highlighting other relevant research, but rather – by mainly understanding the current articles as a part of an evolving field of research as seen through the lens of the two periods seen as a whole.

The main initial finding in the present study is that all but one of studies identified in this follow-up systematic literature review at least in-part support Constantino et al.'s (2011) conclusion of there being a significant association between outcome expectancies and treatment outcomes. However, there are some key considerations that might limit our ability to translate Constantino et al.'s finding into clinical practice or indeed to genialize the results, as explored in detail below.

The first consideration is one that can be found in the systematic literature review the present paper proceeds from (Bolstad, 2017). Bolstad presented four studies that demonstrated a significant association between outcome expectancies and clinical outcomes (Beard et al., 2016; Leibert & Dunne-Bryant, 2015, Snippe et al., 2015; Hedman et al., 2012), three studies that partially demonstrated said association (Delgadillo et al., 2016; Jakubovski & Bloch, 2016; Renaud et al., 2013) and two studies that found no significant association (Barber et al., 2016; LeBeau et al., 2013). Bolstad concluded that these studies did not support Constantino et al.'s

(2011) conclusions. While several of his arguments supporting his conclusion mirrors those outlined in more detail below, his most salient argument was two-fold. Firstly, that the clinical "roles" of expectancies are not only more complex and in-constant than the study designs in question are able to detect and track, the roles also evolve as the treatment progresses. Bolstad's second argument is that the potential roles of expectancies are arguably interacting with other mediators and moderators that are active both before and during treatment than can be sufficiently explored within the remit of the designs chosen in the studies. While both observations are in order and could be applied to several of the studies presented in the present study as well, it is still hard to levy it at any one of the studies as they did not set out to explore the role of outcome expectancies as such a holistic and dynamic construct. The saliency of the observation remains though, asking "if" the association is present is perhaps neither as illuminating regarding clinical insights nor as in enabling in the development of clinical procedures and processes as asking "how" the association is present.

Furthermore, a key objective when deciding to pursue this follow-up systematic literature review, was that it enabled comparing research on the topic across time. What arguably becomes clear is that the research in the present study is a bit different in its focus than in the previous systematic literature review. There is a noticeable shift from most of the articles having a pre-post design and exploring if the association between outcome expectancies and clinical outcomes is present, to the articles in the present study employing randomized control trials designs testing how the association transpires. Several of the articles utilizing a pre-post design in the present study are also arguably more geared toward focusing on the "how" over the "if".

As such, what is arguably the main takeaway from this follow-up systematic literature review is not the results in the studies in isolation or across the current period, but that they when

seen in a broader research effort across both Bolstad's systematic literature review and the present follow-up systematic literature review is that it indicates a gradual re-focusing or shift in direction. Only Barber et al. (2016) article was clearly aimed at answering a "how"-type question in Bolstad's (2017) systematic literature review, while the other articles decidedly focused on it an outcome expectancy – clinical outcome association was present. In the present study the articles from both Chambless et al. (2017), Gobin et al. (2018), Sauer-Zavala et al. (2018), Strauss et al. (2018), Visla et al. (2018) as well as Beck et al. (2021) all illuminate a "how"-type question, while simultaneously providing an answer to the "if"-type question. This shift in overall focus could be a sign of the research on the topic maturing, in that is focusing less on if the association is present and more one the mechanisms of the association.

There are several other factors across the various studies that support such a shift in focus. A criticism that can be levied against both Constantino et al.'s (2011) and Bolstad's (2017) research is that a large part of the samples from which the conclusions are drawn arguably exhibit an inherent positivity bias. The research only includes the expectation of those who not only seek, but also completes the treatment in question. While this is not necessarily problematic in studies that seek to explore if there are certain mechanisms that are in play with a limited sample, such as for instance those who complete psychotherapeutic treatments, one can still question whether any results from such studies can be generalized to be applicable to the general population at large.

Furthermore, the sample is arguably even less representative than one potentially would assume as most of the studies that have been included have taken place within the US' pay-to-play public health care system. As such, the sample is likely further limited to consisting of a limited subset of patients that due to socio-economic factors such as possessing a qualifying

health insurance or that posses the economic means to enter into therapy chooses to do so. This potentially excludes patients who would like to seek treatment regardless of their outcome expectancies, but do not possess the financial opportunities to enter into treatment. The sample also fails to address those who do not seek treatment due to for instance stigma, or due to very low outcome expectations. Less focus is also provided those with negative outcome expectancies, if they seek help or not. Admittedly, exploring and determining how especially positive outcome expectations potentially enables positive clinical outcomes can potentially provide insights that can be put to good clinical application in spite of these reservations. For instance, one example of such an occurrence is if said insights can help facilitate and induce more positive outcome expectations in patients that presently do not enter into treatment due to low outcome expectancies or stigma. However, the issue with generalizing the results from a limited potentially biased sample remains as something to be mindful of.

A potential way to address these potential limitations in future research is to strive for it to be designed in such a way as being treatment specific and where the large-sample statistical studies are preceded by qualitative studies whose goal is to help explore potential mechanisms at play at a smaller scale initially. Any insights provided by the smaller sample or qualitative studies can then potentially be integrated when designing follow-on large-sample studies using statistical analysis to test if these factors are in play. An additional step could be to also ensure that future studies take place in countries that have universal health care systems, potentially enabling what is arguably more representative samples.

A need for more detailed understanding of the samples going forward?

There could be a need for a more thorough understanding of the characteristics of the participants in future studies, especially if the goal of the research changes from if-type studies to how-type studies exploring mechanisms.

For instance, Gobin et al. (2018) sought to identify factors that may explain why female veterans showed poorer treatment response than civilian women when treated for post-traumatic stress disorder (PTSD). The all-female study the explored had a relatively modest sample at 126 participants, 21 veterans and 105 civilians. Their mean age was 42.1 years old and 47.3 years old respectively. Gobin et al. also provided information on the breakdown regarding the ethnicity as well as information on any previous or ongoing comorbid depressive disorder of the participants. They also compared the two groups based on five variables as presented previously in this paper. However, Gobin et al. did not provide any information on the level of education, income, relationship status or categorization of what type of trauma they had suffered for their sample.

Beck et al. (2021) article on the other hand, provided in their study on the same disorder what is arguably the most detailed breakdown on the characteristics of participants of the eight studies that are included in this follow-up systematic literature review. While their study was all male and consisted exclusively of veterans, they were also able ensure that the two groups that were compared were relatively matched for sample size, for ethnic composition, for level of education, for income, for relationship status and for having combat trauma.

Understanding why Beck et al. arguably have a greater emphasis on the composition of their samples than Gobin et al., likely has to do with what they sought to explore respectively. Beck et al. sought to explore the trajectories of change during Group Cognitive-Behavioral Treatment and Group Present Centered Therapy explored across the five Post-Traumatic Stress

Disorder (PTSD) symptom clusters as defined by DSM-5 (American Psychiatric Association, 2013). Gobin et al. sought to identify factors that may explain why female veterans showed poorer treatment response than civilian women. If one was to simply these two objectives into a single word, then "how" and "if" might arguably apply. As such, one could argue that the more intricate the mechanism being explored is then the more one must ensure that the sample is representative, either through carefully tracking its composition and ensuring that they both are representative, or by having a large enough sample that its characteristics statistically averages out during the statistical analysis.

To understand the importance of this one need only delve into the term "veteran" as well as the disorder in question, PTSD. Veterans are far from a homogenous group. For instance, both Beck et al.'s and Gobin et al.'s studies drew their samples from the armed forces in the United States. If you are an officer or an enlisted within the US armed forces, or if your respective ranks within those two hierarchies is high or low, will likely not only greatly correlate to your cognitive abilities, your level of education, your socio-economic status, your level of experience and to what type combat roles you will probably be exposed to, it likely also greatly predict outside factors such as outside expectancies on your role that might increase the stigma in play before seeking help for mental illness or your desire to seek help[6] (Donahue (Director), 2015). Also, if one also factors in Ramage et al. (2015) research on differences between danger[7]- and non-danger-based[8] traumas, meaning that how you can be traumatized can differ greatly, one begins to understand that veterans with PTSD is no homogenous group. Keeping track of all of these variables could be important when determining if it is within-treatment clinical variables,

---

[6] Seeking help can effectively end your career, either by causing your security clearance to be revoked or your aptitude for more demanding positions to be put into question
[7] Exposure to the threat of death or actual threatened serious injury for yourself or others (Ramage et al.,2015)
[8] Exposure to grotesque or haunting images, sounds, or smells of dead or severely injured humans or animals, traumatic loss, moral injury by self or others (Ramage et al.,2015)

such as outcome expectancies, that affects how the clinical outcomes might turn out to be, or if it is factors that are exogeneous in nature that could be moderating or mediating what transpires in treatment. A key question could thus be, as the mechanisms that are explored grow in complexity, would our understanding of the samples also need to grow more detailed?

A need for a standardized outcome expectancy metric procedure?

How the studies determine the outcome expectancy metric is also somewhat challenging for the overall research effort into outcome expectancies when one looks at the studies in aggregate. For instance, the studies in the present follow-up systematic literature review did not conduct the screening consistently with regards to how or when they conducted the screening. The participants in Terock et al.'s (2017) study were screened using the Questionnaire for the Measurement of Psychotherapy Motivation at treatment onset. Beck et al. (2018) administered the Expectancy Rating Questionnaire at the end of the first group session to determine outcome expectancies. Gobin et al. (2018) screened their participants during the second treatment session using Treatment Expectancy Questionnaire. Sauer-Zavala et al. (2018) administered the Credibility/Expectancy Questionnaire after the second session to ensure that the treatment rationale and treatment plan had been understood, while Chambless et al. (2017) used the same questionnaire either before or after the second session. Strauss et al. (2018) also screened their participants after the second session using a study-specific three-question questionnaire. The participants in Turner et al. (2019) study completed the Expectation of Improvement and Suitability of Treatment Form at session six. It could as such be a bit of a stretch to argue that they were all screening for the same construct.

Furthermore, both Bolstad's (2017) systematic literature review and the present study only had two studies each, Beard et al. (2016) and LeBeau et al. (2013) plus Chambless et al.

(2017) and Sauer-Zavala et al. (2018) respectively, that used the same measurement device, coincidentally The Credibility/Expectancy Questionnaire in both cases. Beard et al. administered the questionnaire prior to the second session, whereas LeBeau et al. administered the questionnaire as a part of the pretreatment assessment. As previously mentioned, Sauer-Zavala et al. (2018) and Chambless et al. (2017) administered the questionnaire after the second and either before or after the second session respectively. As such, there is a demonstrable difference in how the same questionnaire was used across the studies. While the variation can be said to be slight between some of studies, it arguably remains something to be mindful of when assessing if they are indeed screening for the same construct.

Another potentially confounding factor is the use of study-specific measuring devices. Bolstad's review had had three articles, Barber et al. (2014), Delgadillo et al. (2016) and Jakubovski & Bloch (2016) that measured the outcome expectancy through a single question specific to their study, and two articles, Renaud et al. (2013) and Snippe et al. (2015), that similarly to an article in the present study, Strauss et al. (2018), used several questions that also were specific to the study in question. This could indicate that just as there needs to be an agreed upon definition of "outcome expectation" and its factors, there is also arguably a need to if not develop then to at least adopt a standardized measuring device, be it through a specific questionnaire or combination of existing questionnaires in given procedure in future studies if one hope to integrate future research into a broader research effort.

The considerable variation in how and when the outcome expectancy metric was determined makes it hard to argue that the various articles were in fact screening for and exploring the same construct. Please note that the said variation is not necessarily to be considered as methodological weaknesses for any one of the studies when the study in question

is seen in isolation. Yet, it is hard to argue that said variation is not a methodological limitation when one seeks to integrate the results and discussion points from the specific studies into a general body of research on the role of outcome expectancies. After all, it is hard to draw conclusions that can be generalized to a wider population if the studies do not in fact research a sufficiently similar construct. Please note that this point is applicable to the broader research effort on outcome expectations too, and not limited to the studies presented in the present paper alone.

A potential needed next step could thus be if not development, then at least seek the adoption of universally applied measurement or measuring device such as a questionnaire or battery of questionnaires administered in pre-defined procedure, to help determine the outcome expectancy at the onset of treatment and to track it as it progresses through treatment.

Methodological strengths and weaknesses in the studies

The studies are all tailor-made and well suited to the examine their respective objectives, when keeping the aforementioned discussion points on overarching methodological concerns in mind. While it potentially would have been even better if the studies all used a randomized control trial design due to its stronger design, the choice of a pre-post study design makes sense given the objectives of the three studies that chose this design. This is especially the case with the study conducted by Visla et al. (2018) with its focus limited to the role of several factors as mediators and moderators for a specific diagnosis, depression, using a specific clinical treatment, cognitive behavioral therapy. One could always do a follow-up study re-using the same data and adding a second treatment form or diagnosis if one wants to establish if their discoveries can be generalized with a greater degree of certainty.

The studies all use statistical methods prudently regarding their respective objectives. Chambless et al. (2017) and Visla et al. (2018) are both examples of studies that adjusted their analytic approach to better fit their data. Chambless et al. (2017) refrained from using the standard approaches to the longitudinal analysis of treatment data and instead used a shared parameters approach (Ten Have et al., 1998) to analyze the change in treatment over time. Their study did not have data missing at random enabling the use of this more specific statistical approach in their analysis. Visla et al. (2018) on the other hand added a number of proven procedures and recommended steps outlined by previous research of a similar nature into their analysis. Visla et al. did this to lessen the potential adverse effects their dataset could have had on the statistical analysis if these steps were not included. The reasoning for the choices that were made are clearly explained in both articles and are both well-reasoned as seen within the limited statistical expertise this analysis is based upon.

Likewise, a potential weakness in the articles could be that all but one article, Terock et al. (2017) at n=518, have relatively small samples ranging from 91 to 175. While there is no universally agreed upon sample size requirements when conducting a randomized controlled trial design studies, several researchers have proposed some guidelines. Sakpal (2010) proposed the minimum requirement of 200 participants which, after a 10% drop out rate, would yield 90 participants in each group. Chan (2010) proposes that the number of participants is a function of design of the study and measures of outcome, providing charts based on standardized effect size and to what power. While pursing insights provided by Sakpal and Chan could have indicated that the samples in the included studies were insufficient, this exploratory pursuit was abandoned due to the studies arguing convincingly for their chosen approaches and the limited statistical expertise the present study would have had to (what would have been to) second guess their choices.

A weakness in Sauer-Zavala et al.'s (2018) and Visla et al.'s (2018) articles is that they both provided no explicit information on their participants other than their number. This is especially odd in the case of Sauer-Zavala et al. as it is clear from their article that they not only possessed information on their participants, but in fact tested for several factors such as age and sex. It thus seems like an unforced error to use some sporting vernacular. Visla et al. limited the information they provide to that the participants are age 18 or above. Both studies would likely have been perceived as methodologically stronger if said information was included.

Future research

The impression that lingers after having explored the results in Constantino et al. (2011, 2018) two meta-analyses on the association between patients' outcome expectation and their post-treatment outcomes, Bolstad's (2017) systematic literature review that followed from Constantino et al. meta-analysis published in 2011 and the results in this follow-up systematic literature review, is that some of the research is providing an answer to the seemingly less interesting question. Even if one leaves to one side the differences in methodology between the studies, the variation in how "outcome expectancy" is operationalized across studies, and the lack of a universally adopted standardized method such as vetted questionnaire to baseline and track expectations before and during treatment, the clinical applicability of some of the studies still is still arguably fairly moderate.

For instance, while both of Constantino et al.'s meta-analyses have demonstrated a significant association between patients' outcome expectation and their post-treatment outcomes, it can still be argued that the data supporting these results potentially is inherently biased as explored in detail below. As such, the results are perhaps not as indicative of there being an association as one could assume. As such, generalizing such results to "prove" that the

association in and of itself is important clinically seems to be a bit of a stretch. Indeed, Constantino et al. points out in the 2018 meta-analysis that: "Future research needs to (a) test strategies that causally enhance patients' outcome expectancy to improve treatment efficacy and (b) illuminate both patient and therapist contributions to patient outcome expectancy to help tailor clinical practice and training." Thus, a more pressing question than *if* there is an association between patients' outcome expectation and their post-treatment outcomes, is perhaps *how* this association can be pro-actively used to help facilitate beneficial outcomes in therapy.

That one should explore how the management of expectation can be used clinically to enhance outcomes is not a novel idea. Some of the studies presented above is also arguable on the onset of providing an answer. Other research not included in this systematic literature review due to it not meeting the selection criteria is also doing so. For instance, Edwards and Fehm (2023) found that both positive therapy expectations and negative therapy expectations for remission had a significant influence on the success of therapy. Their recommendation that one should routinely record therapy expectations at the beginning of therapy to identify obstacles that might impede a beneficial clinical outcome is arguably a starting point in point in providing such answer. Indeed, Chambless et al. (2017) made a point that also support such an approach when they specified that it was "…conceivable that the expectancy measure could reflect not only the treatment rationale but also the patient's initial reaction to the therapist…" as the patient had at least one full session with the therapist before the outcome expectancy measurement was administered. That the outcome expectancy was not measured at the onset of the treatment was far from the norm, as outlined below, yet the aforementioned quote indicates that outcome expectations is not a static construct. Such considerations arguably emphasize the potentially increased clinical utility in re-focusing future research on how management of expectations can

be used proactively to promote positive outcomes, instead of focusing on if there is an association between outcome expectations and clinical outcomes.

An example of a treatment where the management of expectations is the clinical practice is The Bergen 4-day Treatment (B4DT) treating OCD[9]. The B4DT methodology not only routinely measures and makes the expectations of the patients explicit from the onset of the treatment, it also repeatedly explains the treatment's rational, procedures and past results ensuring that its participants both know what is to come as well as implant an understanding in the patients that the treatment works based on research. By asking four questions[10] at the onset of the treatment followed by explaining the treatment's rational as well as providing the clinical results of the treatment as documented through various studies, the treatment seeks to instill a belief or expectancy in the patients that the treatment will work that not only seems to increase the buy-in from the patients, but also potentially increases outcome expectancies. A key function of the four aforementioned questions and continued repetition and foreshadowing of how the treatment will proceed, is also to be able continuously able to address what reservations that the patients might have concerning the efficacy of the treatment, if their outcome expectancies changes or if the treatment is correct from them. The treatment also routinely tracks the commitment of the patients after one has conducted therapist assisted exposure sessions during four-day treatment "week", ensuring that one can address lapses in commitments or if the outcome expectancy is waning. Importantly, the treatment also tracks the outcomes of its patients

---

[9] Please note that the following is based on the author's understanding of the treatment from being a certified therapist in said treatment. It is not on referencing a manual or manuals that are available to the general public as no such manual has been published at the time of writing.

[10] Answer the following on a scale from 0 to 100, where 0 is none/would not recommend and 100% is very/complete/would recommend.1) *How logical does this treatment seem to you? 2) Based on what you know now, how likely is it that you would recommend this treatment to a friend who was severely affected by a disorder of this type? 3) How much faith do you have in your ability to follow and complete the treatment? 4) How much faith do you have in that the treatment will help you with regards to your disorder*? Any reply that is less than 100% is tested for ambiguity, that is – why the reply is less than a 100% to determine if the reason needs be addressed either then and there or later on in the treatment.

ten days (some patients) and three months (all patients) after the conclusion of the treatment, enabling the therapists to re-initiate the treatment in those cases where the desired outcomes that are in-tune with realistic outcome expectancies are not met. While there are other examples of treatments that are monitoring the symptoms, expectancies, working alliance and/or patient commitment to the treatment being conducted through questionnaires or feedback systems such as Norse Mental Health and Substance Use used in Helse Forde in Norway, these methods are more general and more often than not not as directly tied to the ongoing treatment compared to how B4DT has made it an integrated and integral part of the treatment itself.

As such, this paper proposes that future research is needed to more clearly establish how the management of expectations can be utilized to facilitate favorable clinical outcomes. Note that such research could both be general – such as is the case with research in the common factors across treatments but should also be specific to particular treatments such as is the case outlined above regarding B4DT. While several of the studies included in this systematic literature review addresses the role of expectations both as a moderator and mediator for desired outcomes, qualitative research or more limited studies that is treatment specific might also be in order to help refine what factors and sub-factors future, larger-sample research using statistical analysis could be focusing on when doing the statistical analysis on larger-sample datasets.

Limitations in the present study

The present systematic literature review has several limitations. Several of these have to do with the methodology it is based on.

The decision was made early on to adopt the methodology used by Bolstad (2017) in his systematic literature review for the present study. This included using a search strategy that was very close to, but not identical to the one Bolstad used as outlined previously. While the search

strategy worked in that it reduced the number of hits to something that was feasible to explore within the remit of the present thesis, it still arguably exhibits several short comings making the present follow-up systematic literature review's ability to be comprehensive to being in question. For instance, the search strategy's search terms are few and are what can only be described as relatively unrefined. The term "treatment outcome» does not cater to differing phrasings with the same meaning, such as "outcome of treatment". A more nuanced use of search term would likely have been better when seen from a purely methodologically standpoint and could potentially have yielded additional hits.

Furthermore, Bolstad's chose to limit the employed search fields in the PubMed search. Bolstad made this choice solely to limit the number of hits as the original search yielded too many hits with that focused on somatic research and not psychotherapy. This should perhaps not have come as a surprise as PubMed is primarily a database used to publish studies conducted on somatic illnesses or other conditions treated by physicians. The choice of not refining the search parameters further, such as to a larger degree facilitate a search that would exclude most of the non-psychotherapeutic treatments, but rather limit the search field is not ideal methodologically.

Another factor limiting the present study is Bolstad's choice to limit the number of databases to a mere two. While the remit of the current task of writing a thesis for the postgraduate course in the Professional Studies in Psychology at the University in Bergen naturally makes it logical to limit the number of databases, it is still prudent to point out that doing so will more likely than not mean that the systematic literature review is not comprehensive in its scope. While the introduction of exploratory mining, that is using the references or citations from the selected articles or other relevant work such as Constantino et al.'s meta-analyses (2011, 2018) or Bolstad's (2017) systematic literature review to help identify

other potentially relevant studies, could potentially have limited any negative effects of the limited search strategy to some degree, the choice was made early on to abandon this added step in favor of a more stringent adherence to the main parts Bolstad's methodology. This was again due to the remit of the current thesis, also helped by the fact that adding exploratory mining likely would not have ensured a comprehensive search. While limiting the search to two databases and excluding exploratory mining as am avenue to add potentially relevant articles, one can still argue that PsycINFO and PubMed will arguably still represent a broad swath of the available literature on the topic in question. As such, the current systematic literature review could still be of utility based on this limitation.

The present study also has some additional weaknesses regarding the selection criteria. The choice of limiting the articles to being solely being written in English excludes studies that could have been relevant that were written in other languages. Several such articles were identified during the selection process, of which one, Edwards and Fehm (2023), has been mentioned in the discussion due to its relevancy. Likewise, the exclusion of articles that had children, adolescents and family systems as their samples also has arguably limited the present study. Several articles that could potentially have been relevant were identified during the selection phase but were all excluded due to the selection criteria. Future systematic literature reviews can potentially seek to remedy these exclusions by potentially creating wider selection criteria.

A potential weakness is also that the current paper included articles that explored the role of outcome expectancies outside of its direct association with the clinical outcome in addition to the direct association. The emphasis of such alternative associations might have affected the overall analysis of differences between the periods that were examined in Bolstad's (2017)

systematic literature review and that of the current study, yet it was done after careful
deliberation as the discoveries proved nuanced, relevant, and potentially insightful. Please note
that the inclusion of articles was still exclusively done based on the selection criteria.

The final weakness of the present study is its tendency to analyze the significance of the
results from the included articles from a macro-level, instead of engaging with them at a micro-
level such Bolstad (2017) can be said to have done to larger degree. While this choice can be also
be seen as a strength, as re-conducting an existing study for a new period arguably enables us to
explore if there have been any significant changes to what is being researched, it can also
potentially lessen the study's ability to engage with the minutia of the results in its selected
articles.

Conclusion

Seven of the articles that are published after March 2017 that are included in this follow-
up systematic literature review demonstrated an association between outcome expectancies and
the clinical outcomes in at least one of the treatments being explored in the study in question. Of
these, five articles demonstrated a direct association. As such, the initial finding is that all but
one of studies included in this follow-up systematic literature review at least in-part support
Constantino et al.'s (2011) conclusion of there being a significant, if weak, association between
outcome expectancies and treatment outcomes.

The second main result in the present study is that the research on the field is arguably
maturing. When comparing the nature of the articles included in Bolstad's (20017) systematic
literature review with the articles included in this follow-up systematic literature review that is
utilizing nearly the same methodological approach, one can discern a shift from studies primarily
examining if there is an association between outcome expectancies and the clinical outcomes to

currently exploring the mechanisms active in or acting on such a link. As such, the articles included in the present study arguably demonstrate a shift from "if"-type research to "how"-type research.

However, a few notable methodological concerns arguably become apparent as the body of the research arguably shifts to focusing on said mechanisms. The first of these is that there is arguably an increased significance of the need to either adopt or develop a standardized procedure for baselining, keeping track of and potentially managing outcome expectancies as the treatment progresses. The current research varies greatly regarding how and when such measurements are taken, arguably putting into question if they are screening for the same construct. To develop a standardized questionnaire or adopt a standardized battery of existing questionnaires is arguably a logical next step to help mitigate this potential short coming. Secondly, it could be that an increased emphasis is required regarding the characteristics of the participants in future studies to enable research on these potentially more dynamic and nuanced mechanisms. This focus is arguably needed to enable and ensure that future research is exploring variables that are either moderating or mediating pre- within-treatment, and to potentially eliminate interference from exogenous factors. Several ways of potentially combatting such effects from taking place has been proposed, such as having a greater transparency on sample composition and their balance, to increase the sample sizes, to conducting more treatment-specific qualitative research to help inform follow-up larger sample studies using primarily statistical analysis or facilitating more research in countries with universal health care.

References:

Aggarwal, R., & Ranganathan, P. (2019). Study designs: Part 4 - Interventional studies. *Perspectives in clinical research*, 10(3), 137–139. https://doi.org/10.4103/picr.PICR_91_19

Agras, W. S., Walsh, T., Fairburn, C. G., Wilson, G. T., & Kraemer, H. C. (2000). A multicenter comparison of cognitive-behavioral therapy and interpersonal psychotherapy for bulimia nervosa. *Archives of general psychiatry*, *57*(5), 459–466. https://doi.org/10.1001/archpsyc.57.5.459

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Washington, DC: Author.

Barber, J. P., Zilcha-Mano, S., Gallop, R., Barrett, M., McCarthy, K. S., & Dinger, U. (2014). The associations among improvement and alliance expectations, alliance during treatment, and treatment outcome for major depressive disorder. Psychotherapy research : journal of the Society for Psychotherapy Research, 24(3), 257–268. https://doi.org/10.1080/10503307.2013.871080

Başoğlu, M., Lax, T., Kasvikis, Y., & Marks, I. M. (1988). Predictors of improvement in obsessive-compulsive disorder. *Journal of Anxiety Disorders, 2*(4), 299–317. https://doi.org/10.1016/0887-6185(88)90026-6

Beck, J. G., Clapp, J. D., Unger, W., Wattenberg, M., & Sloan, D. M. (2021). Moderators of PTSD symptom change in group cognitive behavioral therapy and group present centered

therapy. Journal of anxiety disorders, 80, 102386.

https://doi.org/10.1016/j.janxdis.2021.102386

Beard, C., Stein, A. T., Hearon, B. A., Lee, J., Hsu, K. J., & Björgvinsson, T. (2016). Predictors of Depression Treatment Response in an Intensive CBT Partial Hospital. *Journal of Clinical Psychology, 72*(4), 297-310. doi:10.1002/jclp.22269

Bolstad, A. (2017). Sammenhengen mellom utfallsforventninger og utfall av psykoterapi: En systematisk litteraturgjennomgang. https://hdl.handle.net/1956/16015

Borkovec, T. D., & Nau, S. D. (1972). Credibility of analogue therapy rationales. *Journal of Behavior Therapy and Experimental Psychiatry, 3*(4), 257–260. https://doi.org/10.1016/0005-7916(72)90045-6

Browne, J., Cather, C., & Mueser, K. (2021). Common Factors in Psychotherapy. *Oxford Research Encyclopedia of Psychology.* https://doi.org/10.1093/acrefore/9780190236557.013.79

Chakraborty, H., & Gu, H. (2009). A Mixed Model Approach for Intent-to-Treat Analysis in Longitudinal Clinical Trials with Missing Values. RTI Press.

Chambless, D. L., Milrod, B., Porter, E., Gallop, R., McCarthy, K. S., Graf, E., Rudden, M., Sharpless, B. A., & Barber, J. P. (2017). Prediction and moderation of improvement in cognitive-behavioral and psychodynamic psychotherapy for panic disorder. Journal of consulting and clinical psychology, 85(8), 803–813. https://doi.org/10.1037/ccp0000224

Chan Y. H. (2003). Randomised controlled trials (RCTs)--sample size: the magic number?. Singapore medical journal, 44(4), 172–174.

Constantino, M. J., Arnkoff, D. B., Glass, C. R., Ametrano, R. M., & Smith, J. Z. (2011). Expectations. *Journal of Clinical Psychology, 67*(2), 184-192. doi:10.1002/jclp.20754

Constantino, M. J., Coyne, A. E., Boswell, J. F., Iles, B. R., & Vîslă, A. (2018). A meta-analysis of the association between patients' early perception of treatment credibility and their posttreatment outcomes. Psychotherapy (Chicago, Ill.), 55(4), 486–495. https://doi.org/10.1037/pst0000168

Delgadillo, J., Moreea, O., & Lutz, W. (2016). Different people respond differently to therapy: A demonstration using patient profiling and risk stratification. *Behaviour Research and Therapy, 79*, 15-22. doi:http://dx.doi.org/10.1016/j.brat.2016.02.003

Devilly, G. J., & Borkovec, T. D. (2000). Psychometric properties of the credibility/expectancy questionnaire. Journal of Behavior Therapy and Experimental Psychiatry, 31(2), 73-86. doi:http://doi.org/10.1016/S0005-7916(00)00012-4

Donahue, T. (Director) (2015) Thank you for your service [Documentary]. Artemis Rising Foundation. Creative Chaos vmg. Sprayregen Family Foundation

Edwards, L., & Fehm, L. (2023). Positive und negative Therapieerwartungen sagen Therapieerfolg vorher [Positive and negative outcome expectancies predict treatment success]. *Zeitschrift für Klinische Psychologie und Psychotherapie: Forschung und Praxis, 52*(2), 98–107. https://doi.org/10.1026/1616-3443/a000706

Freeston, M. H., Ladouceur, R., Gagnon, F., Thibodeau, N., Rhéaume, J., Letarte, H., & Bujold, A. (1997). Cognitive-behavioral treatment of obsessive thoughts: a controlled study. *Journal of consulting and clinical psychology, 65*(3), 405–413. https://doi.org/10.1037//0022-006x.65.3.405

Gobin, R. L., Mackintosh, M. A., Willis, E., Allard, C. B., Kloezeman, K., & Morland, L. A. (2018). Predictors of differential PTSD treatment outcomes between veteran and civilian

women after cognitive processing therapy. Psychological trauma : theory, research, practice and policy, 10(2), 173–182. https://doi.org/10.1037/tra0000266

Goldstein, A. P., & Shipman, W. G. (1961). Patient expectancies, symptom reduction and aspects of the initial psychotherapeutic interview. *Journal of Clinical Psychology, 17*(2), 129-133. doi:10.1002/1097-4679(196104)17:2<129::AID-JCLP2270170207>3.0.CO;2-T

Greer, F. L. (1980). Prognostic Expectations and Outcome of Brief Therapy. *Psychological Reports*, 46(3), 973-974. https://doi.org/10.2466/pr0.1980.46.3.973

Hayes A.F. (2013). Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-based Approach. The Guilford Press.

Hedman, E., Andersson, E., Ljótsson, B., Andersson, G., Andersson, E., Schalling, M., . . . Rück, C. (2012). Clinical and genetic outcome determinants of Internet- and group-based cognitive behavior therapy for social anxiety disorder. *Acta Psychiatrica Scandinavica, 126*(2), 126-136. doi:10.1111/j.1600-0447.2012.01834.x

Jakubovski, E., & Bloch, M. H. (2016). Anxiety Disorder-Specific Predictors of Treatment Outcome in the Coordinated Anxiety Learning and Management (CALM) Trial. *Psychiatric Quarterly, 87*(3), 445-464. doi:10.1007/s11126-015-9399-6

Kraft C., Klauer T., Schneider W. (2015) Erfassung der Psychotherapiemotivation: Revidierte Kurzform des Fragebogens zur Messung der Psychotherapiemotivation. *Psychotherapeut 60*, 315–321

Kirsch, I. (1997). Response expectancy theory and application: A decennial review. *Applied and Preventive Psychology, 6*(2), 69-79. doi:http://dx.doi.org/10.1016/S0962-1849(05)80012-5

LeBeau, R. T., Davies, C. D., Culver, N. C., & Craske, M. G. (2013). Homework Compliance

    Counts in Cognitive-Behavioral Therapy. *Cognitive Behaviour Therapy, 42*(3), 171-179.

    doi:10.1080/16506073.2013.763286

Leibert, T. W., & Dunne-Bryant, A. (2015). Do Common Factors Account for Counseling

    Outcome? *Journal of Counseling & Development, 93*(2), 225-235. doi:10.1002/j.1556-

    6676.2015.00198.x

Mannarini, S., & Rossi, A. (2019). Assessing mental illness stigma: A complex issue. *Frontiers*

    in Psychology, 9, 2722. https://doi.org/10.3389/fpsyg.2018.02722

Noble, L. M., Douglas, B. C., & Newman, S. P. (2001). What do patients expect of psychiatric

    services? A systematic and critical review of empirical studies. *Social Science &*

    Medicine, 52(7), 985-998. doi:http://dx.doi.org/10.1016/S0277-9536(00)00210-0

Oexle, N., Müller, M., Kawohl, W., Xu, Z., Viering, S., Wyss, C., Vetter, S., & Rüsch, N. (2018).

    Self-stigma as a barrier to recovery: a longitudinal study. *European archives of psychiatry*

    and clinical neuroscience, 268(2), 209–212. https://doi.org/10.1007/s00406-017-0773-2

Ogrodniczuk, J. S., & Sochting, I. (2010). Outcome Expectancy Scale. The University of British

    Columbia: Vancouver, BC, Canada.

Renaud, J., Russell, J. J., & Myhr, G. (2013). The association between positive outcome

    expectancies and avoidance in predicting the outcome of cognitive behavioural therapy

    for major depressive disorder. *British Journal of Clinical Psychology, 52*(1), 42-52.

    doi:10.1111/j.2044-8260.2012.02044.x

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D.,

    Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J.,

    Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E.,

    McDonald, S., McGuinness, L. A., … Moher, D. (2021). The PRISMA 2020 statement:

an updated guideline for reporting systematic reviews. *BMJ* (Clinical research ed.), 372, n71. https://doi.org/10.1136/bmj.n71

Pescosolido, B. A., Halpern-Manners, A., Luo, L., & Perry, B. (2021). Trends in Public Stigma of Mental Illness in the US, 1996-2018. *JAMA network open*, *4*(12), e2140202. https://doi.org/10.1001/jamanetworkopen.2021.40202

Porter, E., & Chambless, D. L. (2015). A systematic review of predictors and moderators of improvement in cognitive-behavioral therapy for panic disorder and agoraphobia. *Clinical psychology review, 42*, 179–192. https://doi.org/10.1016/j.cpr.2015.09.004

Ramage, A. E., Litz, B. T., Resick, P. A., Woolsey, M. D., Dondanville, K. A., Young-McCaughan, S., Borah, A. M., Borah, E. V., Peterson, A. L., Fox, P. T., & STRONG STAR Consortium (2016). Regional cerebral glucose metabolism differentiates danger- and non-danger-based traumas in post-traumatic stress disorder. *Social cognitive and affective neuroscience, 11*(2), 234–242. https://doi.org/10.1093/scan/nsv102

Sakpal T. V. (2010). Sample size estimation in clinical trial. *Perspectives in clinical research*, 1(2), 67–69.

Sauer-Zavala, S., Boswell, J. F., Bentley, K. H., Thompson-Hollands, J., Farchione, T. J., & Barlow, D. H. (2018). Expectancies, working alliance, and outcome in transdiagnostic and single diagnosis treatment for anxiety disorders: An investigation of mediation. Cognitive therapy and research, 42(2), 135–145. https://doi.org/10.1007/s10608-017-9855-8

Sibbald, B., & Roland, M. (1998). Understanding controlled trials. Why are randomised controlled trials important?. *BMJ* (Clinical research ed.), 316(7126), 201. https://doi.org/10.1136/bmj.316.7126.201

Snippe, E., Schroevers, M. J., Annika Tovote, K., Sanderman, R., Emmelkamp, P. M. G., & Fleer, J. (2015). Patients' Outcome Expectations Matter in Psychological Interventions for Patients with Diabetes and Comorbid Depressive Symptoms. *Cognitive Therapy and Research, 39*(3), 307-317. doi:10.1007/s10608-014-9667-z

Steketee, G., Siev, J., Fama, J. M., Keshaviah, A., Chosak, A., & Wilhelm, S. (2011). Predictors of treatment outcome in modular cognitive therapy for obsessive-compulsive disorder. *Depression and anxiety, 28*(4), 333–341. https://doi.org/10.1002/da.20785

Strauss, A. Y., Huppert, J. D., Simpson, H. B., & Foa, E. B. (2018). What matters more? Common or specific factors in cognitive behavioral therapy for OCD: Therapeutic alliance and expectations as predictors of treatment outcome. Behaviour research and therapy, 105, 43–51. https://doi.org/10.1016/j.brat.2018.03.007

Sotsky, S. M., Glass, D. R., Shea, M. T., Pilkonis, P. A., Collins, J. F., Elkin, I., . . . Oliveri, M. E. (1991). Patient predictors of response to psychotherapy and pharmacotherapy: findings in the NIMH Treatment of Depression Collaborative Research Program. *American Journal of Psychiatry,* 148(8), 997-1008. doi:10.1176/ajp.148.8.997

Ten Have, T. R., Kunselman, A. R., Pulkstenis, E. P., & Landis, J. R. (1998). Mixed effects logistic regression models for longitudinal binary response data with informative drop-out. *Biometrics*, 54(1), 367–383.

Terock, J., Janowitz, D., Grabe, H. J., Freyberger, H. J., Schneider, W., & Klauer, T. (2017). Alexithymia and Psychotherapeutic Treatment Motivation: Main and Interactional Effects on Treatment Outcome. Psychotherapy and psychosomatics, 86(3), 185–186. https://doi.org/10.1159/000457961

Turner, H., Bryant-Waugh, R., Marshall, E., & Wood, F. (2019). Patient expectations, eating

disorder severity and personality features: Impact on eating pathology in psychological

therapy for eating disorders. *Eating behaviors*, *32*, 85-89.

https://doi.org/10.1016/j.eatbeh.2019.01.002

Vîslă, A., Constantino, M. J., Newkirk, K., Ogrodniczuk, J. S., & Söchting, I. (2018). The

relation between outcome expectation, therapeutic alliance, and outcome among

depressed patients in group cognitive-behavioral therapy. *Psychotherapy Research*, *28*(3),

446–456. https://doi.org/10.1080/10503307.2016.1218089

Wampold, B. E. (2015). How important are the common factors in psychotherapy? An update.

World Psychiatry, 14(3), 270–277. https://doi.org/10.1002/wps.20238.