The state of personalized treatment for anxiety disorders: A systematic review of treatment moderators

Rebecca L. Schneider [a], Joanna J. Arch [a,*], Kate B. Wolitzky-Taylor [b]

[a] Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, CO, USA
[b] Department of Psychiatry and the Behavioral Sciences, University of Southern California, Los Angeles, CA, USA

HIGHLIGHTS

• We synthesize findings examining moderators of treatment outcome in anxiety.
• Limited consistent moderators were identified.
• Statistical and sample size quality ratings were assigned to each study.
• Steps for improving future moderation analyses are proposed.

ARTICLE INFO

ABSTRACT

Introduction: The aim of this review was to synthesize findings for moderators of treatment outcome across adult anxiety disorders, obsessive–compulsive disorder, and posttraumatic stress disorder.
Methods: Twenty-four papers that compared two or more active treatments (at least one of which was a form of cognitive behavioral therapy [CBT]) were identified and organized into five treatment comparison categories (distinct psychotherapy combinations, CBT full package vs. single components, CBT vs. augmented CBT, CBT delivery methods, and CBT vs. pharmacotherapy). Sixty-three distinct baseline moderators were tested across seven categories (symptom severity, comorbid emotional disorders or emotional reactivity, cognitive mainte-nance factors, behavioral maintenance factors, personality traits and disorders, sociodemographic factors, and bi-ological factors).
Results: Few consistent treatment moderators were identified. All studies testing quadratic effects found at least one significant non-linear moderator or predictor effect. In addition, the majority of studies had methodological problems and limitations, demonstrating the need for future methodological improvements.
Conclusion: Limited conclusions can be drawn about how to match anxiety disorder patients to treatment. A strong need to improve the methodological consistency and rigor of treatment moderator studies was identified. A series of recommendations for moderation analyses are proposed in order to strengthen future studies and fa-cilitate replication efforts.

© 2015 Elsevier Ltd. All rights reserved.

Contents

* Corresponding author at: 345 UCB Muenzinger, Boulder, CO 80309-0345, USA. Tel.: +1 303 492 4634; fax: +1 303 492 2967.
E-mail address: Joanna.Arch@colorado.edu (J.J. Arch).

1. Introduction

Over the past three decades, clinical psychology and psychiatry have made impressive gains in developing effective treatments for anxiety disorders. The number of empirically supported treatments (ESTs) for anxiety disorders continues to grow (Chambless & Ollendick, 2001; Chambless et al., 1996, 1998), with cognitive-behavioral therapies (CBT) amassing the broadest research support (Hofmann & Smits, 2008; Norton & Price, 2007; Olatunji, Cisler, & Deacon, 2010). Meta-analyses have demonstrated the superiority of CBT relative to no treatment, placebo, and waitlist controls (Butler, Chapman, Forman, & Beck, 2006; Hofmann & Smits, 2008), as well as to other psychotherapies (Butler et al., 2006; Tolin, 2010). Thus, CBT has been proposed to represent the first line of anxiety disorder treatment (e.g., Arch & Craske, 2009). Despite this advantage, a review of meta-analyses identified CBT response rates of 50% or less for two of the three DSM-IV anxiety disorders for which meta-analyses have been conducted (Hofmann, Asnaani, Vonk, Sawyer, & Fang, 2012). Thus, even within this gold standard, first-line treatment, there is substantial room for improvement.

Personalized medicine – using a person's unique characteristics to tailor treatment – represents one potential solution. With the completion of the Human Genome Project in 2003, medical researchers have been able to examine how an individual's genetic makeup influences disease expression at a molecular level, with the goal of guiding and improving diagnosis and treatment (Ginsburg & Willard, 2009). Treatments for cancer, cardiovascular disease, infectious disease, as well as transplantation medicine have all begun to use personalized medicine (Ginsburg & Willard, 2009). More recently, the call for personalized medicine has been directed towards the mental health field (Insel, 2009; Simon & Perlis, 2010), with the stated goal of improving the efficacy of already established efficacious treatments. Matching people to the best treatment for their particular characteristics, if possible, could increase the effectiveness of that treatment for them, resulting in greater efficacy overall.

Personalized medicine embodies the broader idea that personal characteristics can guide clinicians' treatment decisions to maximize efficacy. In the context of treating anxiety disorders, personalized medicine is helpful to the extent that treatment response is dependent upon measurable patient characteristics. Markers of treatment selection thus need not be limited to markers of molecular-level disease expression. Genetic and neural markers as well as self- and clinician-assessed personality, clinical severity, or sociodemographic indices may each have the potential, theoretically, to guide personalized treatment of anxiety disorders. The latter are not necessarily as precise as the former, but if they can be reliably measured, they are easier and faster to assess, cost-effective, and readily available to mental health practitioners

Identifying baseline characteristics that distinguish people who respond differentially to one treatment versus another – known as treatment moderators (Kraemer, Wilson, Fairburn, & Agras, 2002) – provides an important pathway to moving towards personalized medicine. Treatment moderators represent prescriptive variables that identify which treatment works best for whom and under what conditions (Kraemer et al., 2002). Moderators may be particularly informative if they are relevant to the theoretical mechanisms of how a given treatment is theorized to work. For example, one might hypothesize that socially anxious adults with more pronounced cognitive biases would do better in cognitive therapy than in interpersonal therapy. In contrast, nonspecific predictors of treatment outcome provide prognostic information about how baseline characteristics affect outcome overall, *irrespective of treatment type* (see Fig. 1 for an illustration of the difference between moderators and predictors). Whereas predictors provide important prognostic information about which patients are more likely to respond to treatment, moderators demonstrate how patients' baseline characteristics interact with treatment type to produce different outcomes, yielding the potential for prescriptive recommendations that more directly inform personalized medicine.

Although the majority of work to date has focused on identifying predictors, interest in identifying treatment moderators has grown since the National Institutes of Mental Health's call for a focus on personalized medicine (Insel, 2009). Moderator findings from the adult

and childhood depression and childhood anxiety disorder literatures have been synthesized in reviews (Nilsen, Eisemann, & Kvernmo, 2013; Simon & Perlis, 2010). For the adult anxiety disorder literature, numerous meta-analyses and reviews of treatment *predictors* have been published (Eskildsen, Hougaard, & Rosenberg, 2010; Keeley, Storch, Merlo, & Geffken, 2008; Luborsky, Auerbach, Chandler, Cohen, & Bachrach, 1971; Mennin & Heimberg, 2000; Mululo, de Menezes, Vigne, & Fontenelle, 2012; Olatunji, Davis, Powers, & Smits, 2013; Pampaloni, Bruscoli, & Pallanti, 2004; Solvason, Ernst, & Roth, 2003; Steketee & Shapiro, 1995; Taylor, Abramowitz, & McKay, 2012), however, apart from obsessive–compulsive disorder (Knopp, Knowles, Bee, Lovell, & Bower, 2013), we lack a corresponding synthesis of treatment *moderators*.

This review aims to address this important gap in knowledge on treatment moderators for anxiety disorders. Specifically, we aim to assess and synthesize the published findings that investigate putative treatment moderators with applicability across the current adult anxiety disorders, posttraumatic disorder (PTSD), and obsessive–compulsive disorder (OCD), which is in line with the DSM-IV-TR anxiety disorders. These disorders share common features related to elevated threat sensitivity across subjective, psychophysical and neural levels (see Craske et al., 2009). This transdiagnostic approach reflects the current push to look beyond traditional diagnostic symptom clusters, both in research (e.g., NIMH Research Domain Criteria) and in treatment

I. Predictor
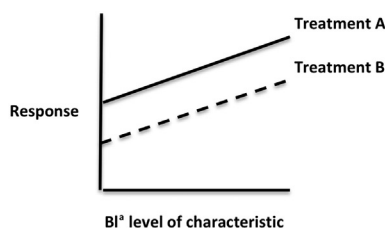
Greater levels of the baseline characteristic predict better response regardless of treatment type.

II. Nonsignificant Result
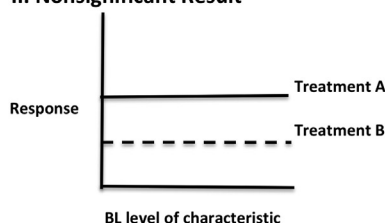
Level of the baseline characteristic does not affect response rate for either treatment.

III. Moderator
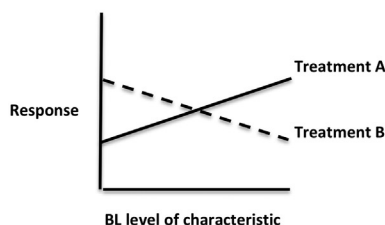
Greater levels of the baseline characteristic predicts *better* response for Treatment A but *worse* response for Treatment B.
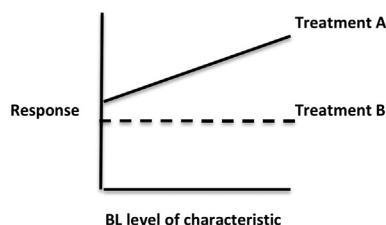
Greater levels of the baseline characteristic predicts better response for Treatment A but does not affect the rate of response for Treatment B.
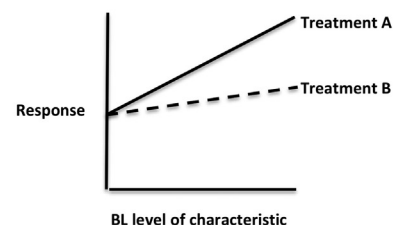
Greater levels of the baseline characteristic predicts better response for both treatments, but to an even greater extent for Treatment A.

[a]BL = baseline; Please refer to Kraemer, 2013 for a similar figure

Fig. 1. Distinguishing treatment predictors and moderators.

(e.g., Unified Protocol, Barlow, Allen, & Choate, 2004). Identifying moderators across the anxiety disorders, rather than examining them within narrow diagnostic categories, improves the clinical utility of this review as well. For example, in primary care settings, patients are commonly given an Anxiety Disorder, Not Otherwise Specified (NOS) diagnosis rather than specific anxiety disorder diagnoses that require conducting a time-consuming diagnostic assessment (Walters, Rait, Griffin, Buszewicz, & Nazareth, 2012). We thus aim for our transdiagnostic moderator review approach to more readily inform actual practice, and will aim to provide direct clinical recommendations based on the findings when possible.

In addition, we focus on treatment studies for which at least one treatment arm consists of the gold standard treatment for anxiety disorders — cognitive behavioral therapy (Arch & Craske, 2009). Far from representing a singular treatment, CBT increasingly represents an umbrella for a range of therapies derived from cognitive or behavioral principles, most of which have shown efficacy for treating anxiety disorders (e.g., Hofmann, Sawyer, Witt, & Oh, 2010; Olatunji et al., 2010). Defining CBT more broadly thus represents an inclusive approach. By focusing the review on treatment studies that include a form of CBT, we aim to elucidate whether particular forms of CBT or competing treatment approaches (pharmacotherapy, third-wave behavior therapies[1], other active psychotherapies) perform better for specific individuals.

Moderator findings can best inform future investigations and clinical practice if they are reliable, rigorous, and well powered. In addition to reviewing findings, we therefore critically examine the statistical and methodological quality of the extant work on anxiety disorder treatment moderators, assessing the extent to which we can expect this work to replicate and to maximally inform questions of personalized medicine. Our approach similarly aims to provide basic recommendations regarding how best to undertake future treatment moderator investigations. To meet this aim, we will identify consistent problems or limitations and outline a step-by-step approach for conducting high quality tests of treatment moderation to encourage researchers to include moderator analyses whenever possible in future treatment studies, as recommended by leading methodological experts (e.g., Kraemer, Frank, & Kupfer, 2006).

In summary, this review aims to:

1. Synthesize (adult) anxiety disorder treatment moderator findings and provide clinical prescriptive recommendations when possible, based on the available literature;
2. Identify knowledge- and methodological-related gaps and limitations in this literature;
3. Propose a series of steps and recommendations for future moderation analyses with the aim of addressing the identified gaps and limitations and improving future efforts to identify treatment moderators.

2. Methods

2.1. Inclusion criteria

We limited our literature search to papers published or in press[2] in English through May 2014. In addition, the following criteria were used to select which studies to include in our review:

1. Participants had to be: (a) 18 years of age or above, and (b) diagnosed with a primary DSM-IV, DSM-III-TR, or DSM-III anxiety disorder. This includes obsessive–compulsive disorder (OCD) and posttraumatic stress disorder (PTSD), given that these disorders had been classified as anxiety disorders until the recent publication

of the DSM-5, and thus were considered anxiety disorders in all research evaluated in this review. We have chosen to include all anxiety disorders classified as such in the DSM-IV since that is the period during which most of the reviewed literature was conducted Studies that failed to conduct a diagnostic interview were excluded (e.g., Zettle, 2003). Although some related reviews have specified a minimum sample size as an inclusion criterion (Nilsen et al., 2013), we chose to evaluate studies regardless of sample size because relatively few studies had met our other inclusion criteria and because evaluating the methodological quality of the extant literature (including sample size) represented a focus of this review.
2. Treatment conditions had to: (a) compare two active treatments, rather than an active treatment versus waitlist or placebo control, in order to examine treatment moderation between two active interventions (thus meeting the definition of treatment moderation, see Kraemer et al., 2002); and (b) include at least one cognitive (CT), behavioral (BT), or cognitive behavioral therapy (CBT) treatment arm. This included complex, multi-component treatment packages such as cognitive therapy and traditional cognitive behavioral therapy (CBT), as well as less complex treatments or treatment components tested alone such as applied relaxation, systematic desensitization, exposure therapy, exposure and response prevention, cognitive restructuring, and cognitive skills training. In that CT, BT, and CBT represent the psychotherapies with the most robust evidence base for treating anxiety disorders (see *Introduction*), we chose to focus on studies that examined moderators of these psychotherapies compared with another variant of CT, BT, or CBT, another distinct psychotherapy treatment, or medication (pharmacotherapy). Therefore, we planned to exclude treatment moderator studies that compared two medications (although we found no such studies) and excluded studies that compared two non-CBT forms of psychotherapy (e.g., compared two types of psychodynamic psychotherapy; Joutsenniemi, Laaksonen, Knekt, Haaramo, & Lindfors, 2012) for anxiety disorders.
3. Putative treatment moderators had to be: (a) variables measured at baseline, prior to treatment initiation; (b) statistically assessed by testing a moderator by treatment condition interaction predicting treatment outcomes (see Baron & Kenny, 1986; Kraemer et al., 2002); and (c) applicable to two or more different DSM-IV, DSM III-TR, or DSM-III anxiety disorders (e.g. disorder severity, sociodemographic characteristics) in order to allow for conclusions to be drawn across multiple anxiety disorders. Based on these criteria, we excluded papers that measured variables after baseline (Burton, Schmertz, Price, Masuda, & Anderson, 2013; Rodebaugh, Heimberg, Schultz, & Blackmore, 2010) or only conducted within-group analyses (Andersson, Carlbring, & Grimlund, 2008; Brody et al., 1998; Cottraux et al., 1990; De Kleine, Hendriks, Kusters, Broekman, & van Minnen, 2012; Hoexter et al., 2013; Nordgreen et al., 2012; Taylor, 2003). We also set out to exclude studies that solely investigated disorder-specific moderators and to report only the transdiagnostic (relevant to multiple anxiety disorders) moderators in studies that examined both disorder-specific (e.g., type of social anxiety disorder, Hedman et al., 2012; dominance of obsessions vs. compulsions, Hohagen et al., 1998; presence of hoarding, Maher et al., 2010) and transdiagnostic moderators.

2.2. Literature search

To identify studies to include in the review, literature searches were conducted using PsycInfo and Pubmed. The search terms "anxiety," "anxiety disorders," "obsessive compulsive disorder or OCD," "post traumatic stress disorder or PTSD," "trauma," "social anxiety disorder or social phobia or SAD," "generalized anxiety disorder or GAD," "specific phobia," and "panic disorder" were paired with the terms "moderator," "moderation," "treatment moderators," "differential prediction," and "neuro predictor," and with the terms "treatment," "treatment

---

[1] Although some may consider third-wave therapies such as acceptance and commitment therapy (ACT) to be forms of CBT, we consider these to represent distinct approaches for the purposes of this review.

[2] We included two papers in press from our research group that were subsequently published.
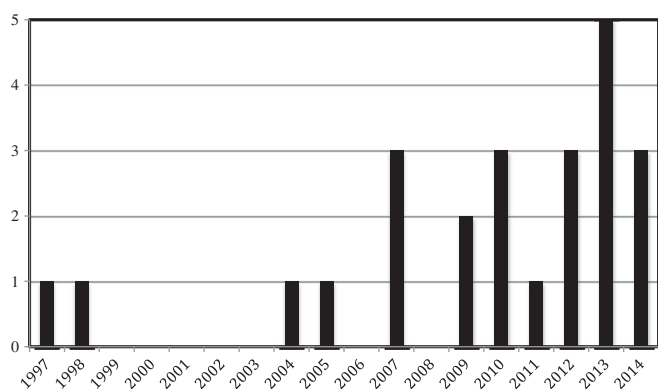
Fig. 2. Number of anxiety disorder moderator papers published or in press through May 2014.

outcome," "outcome," and "treatment response." Searches were conducted with and without including "meta-analysis" and "review" as search terms. Titles from the resulting lists of articles were scanned, and if a title seemed relevant the abstract was read. If a study continued to appear relevant, the results section was skimmed to identify whether moderation analyses were in fact conducted. Since we defined moderation by the statistical analyses that were conducted, it was not necessary for an article to specifically state that it tested for moderation (e.g., Tyrer, Seivewright, Ferguson, Murphy, & Johnson, 1993). Similarly, an article that represented itself as a moderator or differential treatment prediction analysis was excluded if it did not conduct a moderator by condition interaction test (e.g., Hoexter et al., 2013). If moderation analyses were conducted, the article was read in its entirety to identify whether it met our inclusion criteria, as well as to extract relevant information about study design, outcome measures, and moderators. After exhausting all possible combinations of search terms, the reference sections of identified articles were examined for additional relevant studies. A portion of each relevant study title identified from a reference section was entered into the database and both the study and the list of search results generated from the title segment were reviewed. The first author conducted the literature search, with regular input from the second and third authors.

2.3. Search results and study categorization

The initial search yielded 832 unique results in PsycInfo and 480 in Pubmed. After identifying duplicates across search engines and applying our inclusion criteria, a total of 24 published papers based on 21 original studies were identified (see Table 1 in appendix). As presented in Fig. 2, most of these 24 papers were published recently, suggesting growing interest in treatment moderation for adult anxiety disorders. These papers assessed treatment moderators across 19 different combinations of treatment conditions, which we organized into the following five overarching categories: 8/24 papers compared types of CBT[3] to a distinct psychotherapy, 6/24 compared CBT or CBT augmented with medication to medication alone, 4/24 compared CBT to CBT with an added component (e.g., CBT vs. CBT with a motivational interviewing component; Westra, Arkowitz, & Dozois, 2009), 3/24 compared CBT to individual components of CBT, and 3/24 compared different delivery methods of CBT.

Sixty-three treatment moderators were tested across the 24 papers, which we classified into the following seven content categories: 11/63 putative moderators represented *personality traits and disorders*, including DSM-IV personality disorders, five-factor model traits, and motivation; 11/63 represented *symptom levels or baseline severity*, encompassing overall severity of the disorder or a primary symptom of the disorder,

---

[3] CBT herein refers to a type of CBT from among the variants listed in the inclusion criteria.

duration, functional impairments, psychiatric treatment history, and general psychopathology (e.g., severity measured across a broad range of symptoms using the Symptom Checklist-90); 10/63 represented *social and demographic contextual factors*, encompassing traditional demographics (age, sex, race, etc.) and contextual factors such as personal loss, social support, and degree of acculturation; 10/63 represented *biological and neuropsychological* moderators, defined as neural, genetic, and psychophysiological variables, including current use of pharmacotherapy; 9/63 represented *cognitive maintenance factors*, including cognitive misappraisals and biases; 9/63 represented *comorbid emotional disorders or emotional reactivity*, most commonly relating to comorbid mood or anxiety disorders or symptoms; and 3/63 represented *behavioral maintenance factors*, including experiential and behavioral avoidance. Moderators were measured using self-report, clinical interviews, laboratory tasks, neuropsychological tests, genetic tests, and brain scans.

2.4. Our analytic approach

Due to the large variability among the studies in terms of moderator variables, sample types and sizes, and outcome measures, a quantitative meta-analytic approach was not feasible. Rather, studies were systematically organized in accordance with the above criteria, and summarized qualitatively. Predictors were reported only when null moderators were subsequently tested as predictors. In addition, we developed quantitative rating criteria for assessing the methodological rigor of each paper. As Table 1 (see appendix) presents, we assigned one rating for each study's statistical quality and one rating for each study's sample size. The second and third authors independently rated each study's statistical quality, resulting in an interrater reliability ICC of .75. Disagreements were resolved through discussion until consensus was reached. It is important to note that studies receiving the highest quality score do not necessarily meet the highest possible research standards; rather, this score only indicates that the necessary statistical analyses were conducted. Important research design characteristics (e.g., sufficient power, no median splits or arbitrary categories for continuous variables) are not captured in this score.

For statistical quality, we used the following ratings (higher ratings indicate higher quality):

4 The study conducted moderator by condition interaction tests, between-group post hoc tests using the entire sample (e.g., conducting analyses for high and low levels of anxiety sensitivity and examining whether level of anxiety sensitivity affects outcome differentially between two treatments). Preferably though not required because technically moderation is a between-groups test, the study also conducted within-group post hoc tests (e.g., using the same approach and examining whether there is a significant effect of anxiety sensitivity on outcome within each treatment).

3 The study conducted moderator by condition interaction tests and post hoc tests to explore the nature of the interaction, but either conducted them using subgroup analyses without using the full sample (e.g., selecting only participants with comorbid mood disorders and examining whether there was a difference in outcome between the two treatments among this subgroup), or conducted within-group post hoc tests only (e.g., selecting one of the two treatments and examining whether there was a difference in outcome between those with and without comorbid mood disorders only within that treatment).

2 The study examined the effect of the moderator between treatment conditions only within a single level of the moderator (e.g., selecting men and examining whether CBT outperformed medication among men). This rating is similar to the rating of 3, but lacks a full-sample moderator by treatment condition interaction test.

1 The study conducted initial moderator by condition interaction tests but did not explore the nature of the interaction with any follow-up tests.

For sample size, our ratings reflect the following codings of average sample size per condition:

6 Sample size of 75 or above; **5** Sample size of 60–74; **4** Sample size of 45–59; **3** Sample size of 30–44; **2** Sample size of 15–29; **1** Sample size of less than 15.

3. Results

Table 1 (see appendix) presents the 24 reviewed moderator papers and their statistical tests, methodological quality, and sample size.

3.1. Distinct psychotherapy combinations

Five intervention studies produced eight papers examining moderators of treatment outcomes when comparing traditional CBT to another psychotherapy or active psychosocial intervention. Three studies (representing five of the eight papers) compared CBT to a mindfulness- or acceptance-based intervention (acceptance and commitment therapy or an adaptation of mindfulness-based stress reduction). One study (representing two papers) compared CBT to a psychodynamic intervention (panic focused psychodynamic psychotherapy), and one compared CBT to a physiology-focused intervention (capnometry-assisted respiratory training).

3.1.1. Symptom severity

Symptom severity was examined as a putative moderator of treatment outcomes in two of the eight papers. In a study comparing (traditional) CBT to acceptance and commitment therapy (ACT) for mixed anxiety disorders, clinician-rated baseline severity neither moderated nor predicted outcomes (Wolitzky-Taylor, Arch, Rosenfield, & Craske, 2012). In contrast, a study comparing group CBT to adapted mindfulness-based stress reduction (adapted MBSR) for mixed anxiety disorders (Arch & Ayers, 2013) found that baseline anxiety disorder severity moderated post-treatment (but not 3-month follow-up) outcomes. Between-group moderator effects showed that CBT nearly outperformed adapted MBSR in patients with mean levels of baseline severity whereas MBSR outperformed CBT in patients with high levels of baseline severity. Within CBT, low to moderate baseline severity predicted lower post-treatment severity whereas higher baseline severity predicted exponentially higher post-treatment severity. Within MBSR, the nonlinear effect was not significant but a linear effect demonstrated that higher baseline severity predicted higher post-treatment severity.

3.1.2. Comorbid emotional disorders and emotional reactivity

Three papers in this category of CBT versus a distinct psychological treatment examined comorbid emotional disorders or emotional reactivity as putative moderators of treatment outcomes. In the two aforementioned studies of mixed anxiety disorders comparing CBT to either ACT (Wolitzky-Taylor et al., 2012) or adapted MBSR (Arch & Ayers, 2013), higher baseline levels of depressive symptoms or comorbid mood disorders predicted (or trended towards predicting) better outcomes within the mindfulness- or acceptance-based intervention, whereas lower baseline depressive symptoms or the lack of comorbid mood disorders predicted better outcomes within CBT. Significant between-group effects showed that CBT outperformed the mindfulness- or acceptance-based intervention at lower or absent levels of depressive symptoms. In contrast, ACT and adapted MBSR outperformed CBT at moderate to high levels of depressive symptoms (at most but not all time points). A study comparing CBT to ACT for social anxiety disorder (Craske et al., 2014) found that comorbid depression marginally predicted poorer fear and avoidance outcomes across the entire sample but did not moderate outcomes differently between conditions.

In the two CBT versus ACT studies, baseline anxiety symptoms or comorbid anxiety disorders neither moderated nor predicted outcomes (Craske et al., 2014; Wolitzky-Taylor et al., 2012). Subjective anxiety levels during hyperventilation and relaxation laboratory tasks (Davies, Niles, Pittig, Arch, & Craske, 2015) and type of principal anxiety disorder (Wolitzky-Taylor et al., 2012) also had no effect on outcomes in these studies.

Only the ACT versus CBT for social anxiety disorder study (Niles, Mesri, Burklund, Lieberman, & Craske, 2013) examined the role of emotional reactivity as a potential moderator of outcomes. Specifically, level of reported affect in response to viewing positive, negative, and neutral images failed to moderate outcomes, although higher negative affect while viewing negative images – but not positive affect while viewing positive images – predicted better symptom outcomes overall, across both conditions (Niles et al., 2013).

In summary, two out of three studies found that the mindfulness- or acceptance-based intervention outperformed CBT for those with moderate to high levels of comorbid depression, whereas CBT outperformed the mindfulness- or acceptance-based interventions for those with lower or no depression. Baseline level of anxiety or anxiety during laboratory tasks did not moderate outcomes in any study, nor did emotional reactivity to valenced images.

3.1.3. Cognitive maintenance factors

Five of the eight papers (representing four studies) in this CBT versus a distinct psychological treatment category examined cognitive maintenance factors. All five papers examined cognitive misappraisals relevant to the targeted anxiety group; two of these also examined perceived control.

3.1.3.1. Cognitive misappraisals.
Anxiety sensitivity significantly moderated outcomes in both the CBT versus ACT (Wolitzky-Taylor et al., 2012) and group CBT versus adapted MBSR (Arch & Ayers, 2013) for mixed anxiety disorders studies, but in different directions. In the study of CBT and ACT (Wolitzky-Taylor et al., 2012), comparing between conditions, CBT outperformed ACT at moderate baseline anxiety sensitivity levels but showed no advantage at lower and higher levels. Within ACT, baseline levels of anxiety sensitivity did not predict outcomes — that is, ACT performed similarly for those at all levels of baseline anxiety sensitivity. Within CBT, those with moderate levels of baseline anxiety sensitivity benefitted most. When comparing group CBT to adapted MBSR (Arch & Ayers, 2013), on the other hand, adapted MBSR outperformed CBT at moderate levels of baseline anxiety sensitivity, but CBT outperformed adapted MBSR at low and high levels (at post only for high levels).

For the treatment of panic disorder, Meuret, Hofmann, and Rosenfield (2010) compared a cognitive skills training (CT), which focused on changing panic-related cognitions, to capnometry-assisted respiratory training (CART), which focused on changing respiratory patterns. Significant moderator by condition interactions demonstrated that higher baseline levels of catastrophic cognitions were associated with greater improvement in CART compared to CT, whereas lower levels of catastrophic cognitions were associated with greater improvement in CT compared to CART.

In comparing CBT and ACT for social anxiety disorder, Craske et al. (2014) found a condition by time interaction for baseline fear of negative evaluation predicting fear and avoidance outcomes, but no simple condition differences were found at any point. However, greater baseline attentional bias – slower disengagement from negative faces (more than half a standard deviation above the sample mean) – predicted marginally greater improvement following CBT than ACT (Niles et al., 2013). Within CBT, greater baseline attentional bias predicted better fear and avoidance outcomes whereas within ACT, baseline attentional bias did not impact outcomes.

3.1.3.2. Perceived control.
The Meuret et al. (2010) study found that CT outperformed CART at low levels of perceived control and CART outperformed CT at very high levels of perceived control (1.75 SD

above the mean). The Craske et al. (2014) study comparing CBT to ACT for social anxiety disorder also evaluated baseline perceived control, and found that higher levels predicted better fear and avoidance outcomes overall but did not moderate outcomes differently by condition.

In summary, CBT outperformed ACT at moderate levels of baseline anxiety sensitivity, but CBT outperformed adapted MBSR at lower and higher levels. CBT also marginally outperformed ACT at higher levels of baseline attentional bias. Perceived control failed to moderate outcomes in comparing CBT versus ACT but did moderate outcomes in comparing CT to CART. Specifically, CT outperformed CART at lower levels of perceived control whereas CART outperformed CT at higher levels of catastrophic cognitions.

3.1.4. Behavioral maintenance factors

All three papers (representing two studies) that assessed baseline avoidance found that it moderated outcomes, but the direction of findings was mixed. In a mixed anxiety disorder sample, ACT outperformed CBT for those exhibiting higher behavioral avoidance during a voluntary hyperventilation task (Davies et al., 2015). When measuring avoidance using the self-reported Acceptance and Action Questionnaire (AAQ), however, the same ACT versus CBT study found that there was a significant time by condition by baseline AAQ^2 (nonlinear) interaction whereby CBT appeared to perform better than ACT among those exhibiting higher experiential avoidance at the 9 month follow-up, although neither the between- or within-group post hoc tests reached statistical significance (Wolitzky-Taylor et al., 2012). In a social anxiety disorder treatment study, higher levels of self-reported experiential avoidance also predicted better outcomes in CBT than ACT, though again only at the 9 month follow-up (Craske et al., 2014). Since avoidance moderated outcomes differently within a single sample depending on how it was measured, it is possible that experiential avoidance moderates ACT and CBT outcomes differently depending on whether it is measured behaviorally or subjectively. If behavioral and self-reported avoidance in fact function as distinct constructs (at least within the confines of these two studies), this would explain these seemingly incompatible results.

3.1.5. Personality traits and disorders

One paper in this category examined personality disorders and two papers examined personality traits as putative moderators. Specifically, a study that compared panic-focused psychodynamic psychotherapy to applied relaxation training (Milrod, Leon, Barber, Markowitz, & Graf, 2007) examined whether the presence of DSM-IV personality disorders moderated treatment outcomes. However, they compared treatment conditions only within specific levels of the moderator, did not report statistical significance for any finding, and failed to perform a moderator by condition omnibus analysis, making conclusions difficult. With those limitations in mind, it appears that the effect size of improvement in panic-focused psychodynamic psychotherapy was greater than in applied relaxation training in general and that this difference was even greater among patients with comorbid cluster C personality disorders than those without this comorbidity (though as noted, those with and without personality disorders were not directly compared). Cluster B personality disorders did not moderate outcomes, though the number of patients with this diagnosis was very small ($ns = 3$–5 per cell).

Regarding personality traits, the comparison of CBT to ACT for social anxiety disorder (Craske et al., 2014) found that extroversion predicted better outcomes overall but did not moderate differently between the treatment conditions. Neuroticism neither moderated nor predicted outcomes. In addition, neuroticism failed to moderate outcomes between CBT and ACT in the mixed anxiety disorder treatment study (Wolitzky-Taylor et al., 2012), but predicted poorer outcomes overall across both conditions.

In conclusion, although the presence of personality *disorders* significantly moderated outcomes in a study comparing panic-focused psychodynamic psychotherapy to applied relaxation training, no

personality *traits* moderated outcomes in two studies comparing CBT to ACT. However, more work is needed in this area as only three out of the eight papers comparing distinct psychotherapy combinations tested personality as a putative moderator.

3.1.6. Social and demographic contextual factors

Three of the eight papers comparing CBT to a distinct alternative intervention investigated sociodemographic moderators. In the two CBT versus ACT studies (Craske et al., 2014; Wolitzky-Taylor et al., 2012), neither age nor race or ethnicity moderated or predicted outcomes. In the mixed anxiety study (Wolitzky-Taylor et al., 2012), sex failed to moderate or predict outcomes. In the social anxiety disorder study (Craske et al., 2014), although women did better overall, this finding did not differ by condition.

Within the group experiencing an interpersonal loss in the six months prior to panic disorder onset, panic-focused psychodynamic psychotherapy outperformed applied relaxation, and this difference was even greater within the group not experiencing an interpersonal loss (Klass et al., 2009). However, the omission of an omnibus moderator by condition interaction test across the full sample limits support for the author's claim that interpersonal loss moderated applied relaxation but not psychodynamic outcomes.

In summary, no clear evidence supported social or demographic contextual factors as moderators of treatment outcome in comparing CBT for anxiety disorders to a distinct alternative intervention.

3.1.7. Biological and neuropsychological factors

Only one study in this category (Davies et al., 2015) examined biological factors as moderators of treatment outcomes. Comparing CBT to ACT in a mixed anxiety disorder sample, heart rate during a baseline relaxation exercise moderated outcomes, with lower heart rate predicting better outcomes in ACT than in CBT. Resting heart rate at baseline and heart rate variability during relaxation neither predicted nor moderated outcomes. Both higher heart rate variability at baseline and higher heart rate during hyperventilation recovery predicted worse outcomes overall, but did not moderate outcomes differently by treatment condition. In summary, more work is clearly needed.

3.2. CBT full package versus single components

Three papers examined moderators (five in total) of outcomes following the full CBT package versus a single CBT component.

3.2.1. Symptom severity

In a study comparing full CBT (e.g., cognitive therapy plus exposure), cognitive therapy only, and self-control desensitization for GAD (Newman & Fisher, 2013), the authors report that GAD duration moderated treatment outcomes, although no post-hoc statistical tests were reported. Specifically, those with longer duration of GAD showed greater reliable change from the single component treatments whereas those with shorter duration showed greater reliable change from full CBT.

3.2.2. Comorbid emotional disorders and emotional reactivity

The presence of comorbid anxiety disorders (GAD or panic disorder) moderated OCD outcomes in a study comparing exposure and response prevention (ERP) plus relaxation versus ERP plus cognitive therapy, though moderation was limited to the intent-to-treat sample (Hansen, Vogel, Stiles, & Gunnar Götestam, 2007). Specifically, in the ERP plus relaxation condition, those with a comorbid anxiety disorder did more poorly than those without this comorbidity, but in the ERP plus cognitive therapy condition, those with a comorbid anxiety disorder did somewhat better than those without this comorbidity. *Personality traits and disorders.* In the aforementioned study comparing ERP plus CT to ERP plus relaxation for the treatment of OCD (Hansen et al., 2007), among treatment completers, patients with comorbid cluster C personality disorders responded better to ERP plus relaxation than to ERP plus

CT. Further, within the ERP plus relaxation condition, comorbid cluster C patients outperformed those who lacked this comorbidity. The presence of comorbid cluster A and B personality disorders did not moderate outcomes between the two therapies, although they predicted worse outcomes across both therapies, at least in the intent-to-treat sample.

3.2.3. Social and demographic contextual factors

Only one study in this category (Thrasher, Power, Morant, Marks, & Dalgleish, 2010) examined social and demographic contextual factors. This study, which compared prolonged exposure (PE), cognitive restructuring, their combination, and relaxation for the treatment of PTSD, found a significant between-group effect whereby greater social support at baseline predicted better outcomes within the three active conditions relative to the relaxation condition, but did not differentially predict among the three active conditions (Thrasher et al., 2010).

3.2.4. Untested categories

No studies examined cognitive maintenance factors, behavioral maintenance factors, or biological and neuropsychological factors as moderators of treatment outcomes for studies comparing CBT to its individual components.

3.3. CBT versus augmented CBT

Four studies investigated moderators of treatment outcomes (17 moderators total) in CBT versus CBT plus an added active component.

3.3.1. Symptom severity

Two papers compared D-cycloserine (DCS)-enhanced versus placebo-enhanced treatments. Baseline severity did not moderate outcomes in either the study adding DCS to prolonged exposure (PE) for PTSD (De Kleine, Hendriks, Smits, Broekman, & van Minnen, 2014) or the study adding DCS to CBT for social anxiety disorder (Smits et al., 2013). However, in the latter study, higher baseline severity predicted a faster rate of improvement across conditions (i.e. steeper symptom decline slope), but not lower post-treatment scores.

Worry severity, on the other hand, moderated outcomes in a study comparing CBT to CBT plus pre-treatment motivational interviewing for GAD (MI-CBT; Westra et al., 2009). Specifically, those with high levels of baseline worry severity improved more in MI-CBT than in CBT alone, whereas no differences emerged among those with low to moderate levels.

3.3.2. Comorbid emotional disorders and emotional reactivity

In both studies comparing CBT or PE to a DCS- or placebo-enhanced version of each therapy, neither depressive symptoms nor presence (De Kleine et al., 2014) or number (Smits et al., 2013) of comorbid Axis I disorders moderated or predicted outcomes. *Personality traits and disorders.* Three studies in this category examined personality traits as putative moderators. In the first study (Westra et al., 2009), participants with low levels of intrinsic motivation responded better to MI-CBT than to CBT alone, whereas those with high motivation responded equally well to both conditions. Although both studies comparing DCS- to placebo-enhanced treatments found level of conscientiousness to be a moderator, Smits et al. (2013) found that *lower* levels of conscientiousness predicted better outcomes in DCS-enhanced than placebo pill-enhanced CBT for social anxiety disorder whereas De Kleine et al. (2014) found that *higher* levels of conscientiousness predicted better outcomes in DCS-enhanced than placebo pill-enhanced PE for PTSD. Findings were also mixed for agreeableness and extraversion traits. Higher levels of agreeableness predicted better outcomes in DCS-enhanced than placebo-enhanced treatment in the social anxiety disorder study (Smits et al., 2013) but not in the PTSD study (De Kleine et al., 2014), whereas lower levels of extraversion predicted better outcomes in DCS-enhanced treatment in the PTSD study (De Kleine et al., 2014) but not in the social anxiety disorder study (Smits et al., 2013). Higher

levels of neuroticism failed to moderate outcomes in both studies, but predicted worse outcomes across conditions in the Smits et al. (2013) study. Openness neither moderated nor predicted outcomes in either study.

In summary, although it appears that the effectiveness of D-cycloserine as an adjunct to cognitive or behavioral therapies may be influenced by certain personality traits, the contradictory nature of the significant findings precludes solid conclusions until further research can clarify the nature of this relationship, or whether a relationship consistently exists (as opposed to being the product of a specific sample or random error).

3.3.3. Social and demographic contextual factors

Six sociodemographic variables were tested as putative moderators in comparisons of CBT versus augmented CBT; only one was significant. In a one-session exposure treatment for specific phobia (primarily spiders), Pan, Huey, and Hernandez (2011) found that a culturally adapted version of the exposure was more beneficial for Asian-Americans with low levels of acculturation than a non-culturally adapted version, but that outcomes did not differ for those with high acculturation.

In both DCS- versus placebo pill-enhanced psychotherapy studies (De Kleine et al., 2014; Smits et al., 2013), marital or cohabitation status, age, and education level failed to moderate outcomes. However, cohabitation predicted better overall outcomes across conditions in the Smits et al. (2013) study and higher education level predicted better overall outcomes in the De Kleine et al. (2014) study. African-American status (Smits et al., 2013) and sex (De Kleine et al., 2014) also failed to moderate outcomes, although the former predicted better outcomes across conditions.

3.3.4. Biological and neuropsychological factors

Current antidepressant use failed to moderate or predict outcomes in the study comparing DCS- versus placebo pill-enhanced prolonged exposure for PTSD (De Kleine et al., 2014).

3.3.5. Untested categories

No studies examined cognitive maintenance factors or behavioral maintenance factors as moderators of treatment outcomes in studies comparing CBT to augmented CBT.

3.4. CBT delivery methods

Three papers examined moderators of treatment outcomes (testing 27 moderators total) when comparing different CBT delivery formats: Internet-based versus live group CBT (El Alaoui et al., 2013; Hedman et al., 2012) and brief versus standard-length CBT (Dow et al., 2007)

3.4.1. Symptom severity

A study comparing a brief, six-week version of CBT (brief CBT) to a standard, 12-week version (standard CBT) for panic disorder demonstrated that those with more severe or enduring panic disorder benefited more from standard CBT whereas those with less severe and enduring panic disorder benefited similarly from both treatments, although no within- or between-group post hoc tests were reported (Dow et al., 2007). Similarly, those with poorer mental well-being improved more in standard CBT whereas those with better mental well-being improved similarly in both treatments. In contrast, baseline frequency of panic attacks did not moderate outcomes. In summary, those with greater baseline severity tended to respond better to full length than brief CBT whereas those with less baseline severity responded similarly to both.

In the study comparing Internet-based CBT (ICBT) with live group CBT (CBGT) for panic disorder (El Alaoui et al., 2013), those with more domestic functional impairment and those with an earlier onset of symptoms improved more from ICBT than from CBGT, although duration failed to moderate or predict outcomes. However, when specifically

measuring baseline severity, both this study and a study comparing ICBT to CBGT for social anxiety disorder (Hedman et al., 2012) found that higher severity predicted worse outcomes overall but did not moderate outcomes differently between conditions. Lower quality of life in the social anxiety disorder study (Hedman et al., 2012) and greater self-reported work impairment in the panic disorder study (El Alaoui et al., 2013) also predicted worse outcomes overall but did not moderate outcomes differently between conditions. Social impairment failed to moderate or predict outcomes (El Alaoui et al., 2013). Since various indices of baseline severity only moderated outcomes 2/8 times when comparing ICBT with CBGT, it appears that more severe patients do not necessarily benefit more from in-person CBT than from Internet-based CBT.

3.4.2. Comorbid emotional disorders and emotional reactivity

In the study comparing ICBT with CBGT for social anxiety disorder (Hedman et al., 2012), comorbid mood disorder moderated outcomes. Specifically, the absence of a comorbid mood disorder predicted better follow-up outcomes in ICBT but not in CBGT, although no post-hoc analyses were conducted. Depression *level*, however, failed to moderate outcomes in this study (though higher levels predicted worse outcomes overall) as well as in the study comparing ICBT to CBGT for panic disorder (El Alaoui et al., 2013) and in the study comparing brief to standard CBT for panic disorder (Dow et al., 2007). Comorbid agoraphobia, social anxiety disorder, and specific phobia also failed to moderate outcomes when comparing ICBT to CBGT for panic disorder (Dow et al., 2007).

In the study comparing ICBT and CBGT for social anxiety disorder, lower general anxiety predicted better follow-up outcomes in ICBT but did not predict outcomes in CBGT (Hedman et al., 2012). Again, however, condition differences were not directly tested.

In summary, although baseline comorbidity moderated outcomes in Internet-based versus live group CBT for social anxiety disorder, baseline comorbidity in both studies for panic disorder and depression levels in all three studies failed to moderate outcomes.

3.4.3. Cognitive maintenance factors

Among panic disorder patients with more severe agoraphobic cognitions, standard CBT outperformed brief CBT whereas the treatments did not differentially perform among those with less severe agoraphobic cognitions. However, conclusions were limitedly based on a moderator by condition omnibus test, and post-hoc tests were not conducted (Dow et al., 2007).

Higher levels of anxiety sensitivity significantly moderated outcomes in the initial regression analysis when comparing ICBT to CBGT for panic disorder (El Alaoui et al., 2013). Specifically, ICBT outperformed CBGT among patients with high anxiety sensitivity, whereas there was no difference among those low in anxiety sensitivity. In a final model that included all moderators, however, higher anxiety sensitivity predicted better outcomes across both conditions but failed to moderate outcomes between conditions.

3.4.4. Behavioral maintenance factors

The severity of agoraphobic avoidance did not moderate outcomes in comparing brief and standard CBT for panic disorder (Dow et al., 2007).

3.4.5. Personality traits and disorders

High levels of adventure seeking and impulsiveness predicted better follow-up outcomes in group CBT but not Internet-based CBT for social anxiety disorder, although it does not appear that within- or between-group follow-up analyses were conducted (Hedman et al., 2012).

3.4.6. Social and demographic contextual factors

Hedman et al. (2012) tested whether computer skill (relevant because of the ICBT arm) moderated outcomes but found no evidence for this. Hedman and colleagues also found that employment status

(i.e. working full time), educational level (i.e. attending college), and having children failed to moderate outcomes, although they predicted better outcomes overall. Age, sex, and employment status also failed to moderate or predict outcomes when comparing ICBT to CBGT for panic disorder (El Alaoui et al., 2013). Thus, none of these six sociodemographic factors were found to differentially predict outcomes in comparing Internet-based and live group CBT.

3.4.7. Biological and neuropsychological factors

The gene polymorphisms 5-HTTLPR, COMTval58met, and BDNFval66met neither moderated nor predicted outcomes for the treatment of social anxiety disorder (Hedman et al., 2012). Current psychotropic medication use neither moderated nor predicted outcomes for the treatment of panic disorder (El Alaoui et al., 2013).

3.5. CBT versus pharmacotherapy

Six studies examined moderators of treatment outcomes (19 moderators in total) in comparing CBT to pharmacotherapy interventions.

3.5.1. Symptom severity

In comparing CBT to pharmacotherapy, three treatment studies for OCD examined baseline severity as a moderator, resulting in mixed findings. The first study compared the addition of ERP versus stress management training to SSRI medication for OCD and found that higher baseline severity predicted worse outcomes within the stress management training plus SSRI condition, but had no affect within the ERP plus SSRI condition (Maher et al., 2010). The second study compared group CBT (CBGT) to the SSRI fluoxetine, and found that greater baseline severity predicted a better response overall, but did not moderate outcomes differently between treatment conditions (D'Alcante et al., 2012). The third study, however, compared psychotherapy (cognitive therapy or ERP) to combination treatment (fluvoxamine with CT or ERP) and found that higher baseline severity did not moderate outcomes between conditions but predicted worse outcomes overall (de Haan et al., 1997). General levels of psychopathology did not moderate or predict outcomes in this study, however.

Neither level of functioning, quality of life, nor past SSRI treatment moderated outcomes in the first study comparing SSRI plus stress management training versus SSRI plus ERP treatment, but lower quality of life and greater number of SSRI trials predicted worse outcomes overall across these conditions (Maher et al., 2010). Duration of OCD neither moderated nor predicted outcomes in the two studies that tested it (de Haan et al., 1997; Maher et al., 2010). Similarly, age of onset did not moderate or predict outcomes (de Haan et al., 1997).

In sum, baseline severity only moderated outcomes when all treatment arms included medication. In the three studies comparing CBT to pharmacotherapy outcomes, baseline severity moderated outcomes in only 1 of 8 instances, suggesting that it generally does not differentially predict CBT versus pharmacotherapy outcomes.

3.5.2. Comorbid emotional disorders and emotional reactivity

Whereas none of the three OCD studies discussed immediately above (D'Alcante et al., 2012; de Haan et al., 1997; Maher et al., 2010) found that level of depressive symptoms moderated or predicted outcomes, baseline depressive symptoms interacted with treatment type to moderate outcomes in a fourth OCD study (Hohagen et al., 1998). Specifically, post-hoc tests demonstrated that for those receiving behavioral therapy plus fluvoxamine, baseline depression levels did not affect outcomes. However, for those receiving behavioral therapy plus placebo, baseline depression levels did affect outcomes such that higher depression predicted worse outcomes (Hohagen et al., 1998).

Tyrer et al. (1993) also found a moderating effect of depression level in a study comparing CBT, medication (dothiepin, diazepam, or placebo), and self-help in the treatment of panic disorder, GAD, or dysthymia. Although statistics for post-hoc within-group tests were not presented,

it appeared that within the self-help condition, patients with greater depressive symptoms improved more than those with less severe depressive symptoms.

In summary with regard to the four CBT versus pharmacotherapy studies for OCD, when all treatment arms of a study included medication, level of depression did not affect outcomes. However, when only one treatment arm included medication, findings for depression as a moderator of OCD outcomes were mixed.

In addition, general anxiety symptoms failed to moderate outcomes in the study comparing CBGT to fluoxetine (D'Alcante et al., 2012).

3.5.3. Cognitive maintenance factors

In comparing CBT to imipramine for the treatment of panic disorder (Hicks et al., 2005), stronger fear of social catastrophe did not moderate outcomes but predicted worse outcomes across both treatments. Physical and mental catastrophe fears neither moderated nor predicted panic disorder outcomes.

Level of insight into the validity of the obsessional cognitions neither moderated nor predicted outcomes in the study comparing SSRI plus stress management training versus SSRI plus ERP for OCD treatment (Maher et al., 2010).

3.5.4. Personality traits and disorders

Number of comorbid personality disorders failed to moderate or predict outcomes when comparing two combination therapies (Maher et al., 2010). However, in the study comparing monotherapies of medication and CBT for GAD, panic disorder, or dysthymia, patients with a comorbid personality disorder improved more with medication than with CBT, although both were superior to self-help (Tyrer et al., 1993). When comparing psychotherapy alone (CT or ERP) to combination therapy for OCD, presence of personality disorder did not moderate outcomes and only predicted worse outcomes when general level of psychopathology was not statistically controlled (de Haan et al., 1997).

Level of motivation did not moderate outcomes, but higher motivation predicted better response overall in the study comparing psychotherapy alone to combination therapy for OCD (de Haan et al., 1997).

In summary, only in the study comparing medication and CBT as monotherapies found that comorbid personality disorders moderated outcomes.

3.5.4. Social and demographic contextual factors

Of the five sociodemographic factors examined in comparing a CBT-related psychotherapy to pharmacotherapy, four were examined across multiple studies. Returning to the OCD study comparing SSRI treatment plus ERP or Stress Management Training (Maher et al., 2010), sex moderated the relationship between treatment type and outcome. Specifically, although men and women receiving ERP both showed lower post-treatment OCD severity than men and women receiving stress management training, the benefit of ERP over stress management training was significantly larger for men than women. In contrast, in the OCD study comparing fluoxetine to CBGT (D'Alcante et al., 2012), sex neither moderated nor predicted outcomes.

Employment status and marital status did not moderate or predict outcomes in a study comparing combination treatments (Maher et al., 2010) or in a study comparing psychotherapy alone to combination treatment for OCD, (de Haan et al., 1997). Years of education failed to impact outcomes either when comparing CBGT to fluoxetine for OCD (D'Alcante et al., 2012). Although age did not impact outcomes when comparing combination therapies for OCD (Maher et al., 2010), older adults responded better to fluoxetine and younger adults responded better to CBGT in another OCD study (D'Alcante et al., 2012).

In summary, sex and age moderated outcomes one of the two times each was tested. Previous treatment, socioeconomic status, employment status, marital status, and education level failed to moderate outcomes in any study.

3.5.5. Biological and neuropsychological factors

One OCD study examined neuropsychological factors as moderators of fluoxetine versus CBGT outcomes, although no post hoc tests were conducted (D'Alcante et al., 2012). Higher verbal IQ predicted better response to both treatments than lower verbal IQ, but to a greater extent in CBGT than in fluoxetine. Verbal memory and learning, measured by the total recall portion of the California Verbal Learning Test (CVLT-II), predicted better response in fluoxetine and, to a lesser extent, in CBGT. Mental flexibility, measured by the number of perseverations in the CVLT-II, predicted better response to CBGT but worse response to fluoxetine. Inhibitory control, measured using the Victoria Stroop Test, both moderated and predicted outcomes. Specifically, fewer errors on the color section was associated with better outcomes, but to a greater extent in CBGT than in fluoxetine. Faster completion of the dots section and slower completion of the words section predicted better outcomes overall, but did not moderate outcomes differently between treatments. Other tests of attention and executive function (Trail Making Test, Wisconsin Card Sorting Test, and Iowa Gambling Task) did not predict or moderate outcomes.

In summary, a single study comparing fluoxetine to CBGT found that mental flexibility, verbal IQ, verbal memory and learning, and one measure of inhibitory control each moderated outcomes, but measures of executive function and a second measure of inhibitory control failed to moderate outcomes.

3.5.6. Untested categories

No studies examined behavioral maintenance factors as moderators of treatment outcomes when comparing combinations of medication and CBT.

3.6. Statistical quality and sample size

As presented in Table 1, nine of the 24 (37.5%) papers received the highest statistical quality score of four. This score does not indicate that the statistical methodology was ideal across all dimensions, but rather indicates that they conducted the necessary between-group post hoc tests using the full sample on at least one moderator variable. Six of the 24 (25%) received a statistical quality score of three, 2/24 (8%) received a score of two, and 7/24 (29%) received the lowest score of 1, representing papers which failed to conduct any statistical analyses beyond the omnibus moderator by condition interaction test. Regarding sample size, one out of 24 (4%) papers drew from an average sample size per condition greater than 75 for at least one analysis, 3/24 (12.5%) had a sample size of 60–74, 4/24 (17%) had a sample size of 45–59, 7/24 (29%) had a sample of 30–44, 9/24 (37.5%) had a sample size of 15–29, and 1/24 (4%) had a sample size less than 15.

4. Discussion

Although increasing attention has been directed towards the identification of treatment moderators, we identified only 24 papers that met our inclusion criteria by testing moderation in the treatment of adult anxiety disorders. To be included, papers needed to statistically test the interaction of a baseline variable with anxiety disorder treatment conditions, with at least one condition comprised of a cognitive or behavioral treatment. To follow, we summarize the overall findings from these 24 papers and their strengths and weaknesses. Informed by the challenges and limitations of the reviewed literature, we put forth a list of recommendations to guide future investigations of treatment moderation.

4.1. Overall findings

Significant variability characterized the anxiety disorder populations, treatments, and moderators tested. Despite this variability across studies, in moderator areas that demonstrated some degree of

replication we have attempted to draw tentative conclusions across populations and treatments (see *Results* for summaries of findings within treatment categories). Combining results in this manner allows us to begin assessing overarching patterns. Due to the small number of studies and the aforementioned variability, however, these conclusions should be considered quite preliminary.

First, basic demographic variables (e.g., sex, age, race) were tested 27 times across nine studies but failed to moderate outcomes 93% of the time. Only sex and age moderated outcomes in single studies, either when comparing combination therapies (sex; Maher et al., 2010) or monotherapies of CBT and medication (age; D'Alcante et al., 2012) for OCD. Although we assume caution in drawing conclusions based on acceptance of a null hypothesis, the evidence overwhelmingly suggests that with the possible exception of OCD, demographic factors have failed to predict whether someone will do better or worse with one type of treatment for anxiety disorders than another.

Drawing conclusions about the moderator status of baseline anxiety disorder severity remains challenging given the complexity of the findings. Over one-third of moderator tests for baseline severity (9/26 tests from 12 studies) were significant, but no clear pattern of findings emerged. More research is needed in order to draw conclusions about baseline severity as a moderator of anxiety disorder treatment outcomes.

One common anecdotal assumption observed in clinical settings is that more severe anxiety disorder cases should be treated with medication. Yet neither of the two studies comparing pharmacotherapy to a cognitive or behavioral therapy found that baseline severity moderated outcomes (D'Alcante et al., 2012; de Haan et al., 1997). Therefore, the extant data, restricted to OCD samples, fails to uphold this assumption.

Findings also were mixed when examining the role of experiential avoidance in ACT versus CBT. Specifically, ACT performed better than CBT at higher levels of behavioral avoidance in one paper (Davies et al., 2015), whereas CBT performed better at higher levels of experiential avoidance in two other papers (Craske et al., 2014; Wolitzky-Taylor et al., 2012), suggesting that the relationship between avoidance level and treatment type is complex. These findings highlight that tests of moderation allow us to assess specific hypotheses about whether a treatment best serves those it targets — in this case, whether ACT, which targets experiential avoidance, in reality represents the best treatment for those high in experiential avoidance. In actuality, however, CBT outperformed ACT at high levels of avoidance two out of the three times it was tested.

The role of anxiety sensitivity was tested in two papers comparing CBT to either an acceptance- (Wolitzky-Taylor et al., 2012) or mindfulness-based (Arch & Ayers, 2013) treatment, again with mixed results. The authors suggest that different mean levels of anxiety sensitivity across the two studies, as well as the different nature of the interventions (ACT vs. adapted MBSR, delivered in individual vs. group formats) likely account for these seemingly divergent findings. However, the fact that CBT performed optimally at moderate (but not high) levels of anxiety sensitivity in one study (Wolitzky-Taylor et al., 2012) again illustrates the counterintuitive finding that treatments designed to target specific pathological processes do not always perform best for those endorsing the highest levels of this pathological process.

In reality, treatment moderation may often reflect non-linear relationships — a notion based on the fact that of the five studies testing quadratic effects (Arch & Ayers, 2013; Craske et al., 2014; Davies et al., 2015; Niles et al., 2013; Wolitzky-Taylor et al., 2012), all found at least one significant quadratic moderator or predictor effect. Although each of these non-linear moderator findings need to be replicated, if they hold up then it would follow that for therapies targeting certain processes, such as ACT targeting experiential avoidance, there is an ideal baseline range associated with maximum treatment benefit. In support, the quadratic effects observed for both experiential avoidance and anxiety sensitivity (Arch & Ayers, 2013; Wolitzky-Taylor et al., 2012) indicate that there may be an optimal level of these variables (i.e., moderate but

not high) in which the treatments can be most effective. For example, it may be possible to be too avoidant of emotions to benefit maximally from ACT or too fearful of physical anxiety symptoms to benefit maximally from CBT. Yet it seems important to endorse at least a moderate level of the targeted pathological symptoms or processes at baseline in order for the therapy targeting them to be relevant. Thus, it is possible that some non-significant or mixed moderator findings stem from the fact that the majority of studies we reviewed examined only linear moderation relationships. Clearly, this idea requires further testing. Moreover, inclusion of more complex statistical modeling that takes into account the levels of more than one putative moderator at a time may provide an even more sophisticated and precise way of identifying *combinations* of patient characteristics (e.g., women with high neuroticism and moderate anxiety sensitivity) that can provide clinical prescriptive recommendations. As we are still in the nascent state of this area of research, this may be a longer-term but important goal.

Of the eight personality *traits* tested, six were assessed across multiple studies. Out of these six, only conscientiousness was found to moderate outcomes both times it was tested, but findings were in opposite directions (De Kleine et al., 2014; Smits et al., 2013). Neuroticism and openness do not appear to serve as treatment moderators. Additional replication efforts are needed to determine what moderating effect – if any – other personality traits have on treatment outcomes.

Personality *disorders*, on the other hand, moderated treatment outcomes 60% of the time (3 out of 5 tests). In two of three studies, simpler treatments such as medication or relaxation training outperformed more psychologically complex treatments such as CBT among those with comorbid personality disorders (Hansen et al., 2007; Tyrer et al., 1993). This finding suggests that those with principal anxiety disorders and co-occurring personality disorders may respond better to less psychologically complex treatments, although further investigation is needed.

Investigations of depression as a moderator were significant in less than forty percent of relevant studies (5/14 tests). Depression levels did not moderate outcomes in two of three studies comparing CBT (with or without placebo) to pharmacotherapy (with or without CBT) for OCD (D'Alcante et al., 2012; de Haan et al., 1997). However, in two of three studies (Arch & Ayers, 2013; Wolitzky-Taylor et al., 2012) higher depression levels predicted better outcomes in an acceptance- or mindfulness-based therapy compared to CBT. If replicated, this finding could inform future efforts at treatment matching.

Relatively few studies examined baseline neural or physiological measures in relation to anxiety disorder treatment response. Of these, even fewer looked at between-group moderators — most examined only within-group predictors (e.g., Brody et al., 1998; Hoexter et al., 2013). Further, aside from two studies examining the impact of current psychotropic medication use (El Alaoui et al., 2013; De Kleine et al., 2014), the three additional studies examining biological moderators of treatment outcomes focused on different biological measures (; D'Alcante et al., 2012; Davies et al., 2015; Hedman et al., 2012). Although there were some significant moderator findings from the physiology and neuropsychology tests, none of the tested genetic markers served as moderators. Thus, putative biological moderators of anxiety disorder treatment outcomes are in particular need of additional work and replication efforts.

4.2. Problems with validity and reliability

Overall, the inconsistent moderator findings across multiple moderator categories likely stems in part from the fact that most studies compared different sets of treatments. However, two studies, conducted by the same research team, compared traditional CBT to ACT in four papers (Craske et al., 2014; Davies et al., 2015; Niles et al., 2013; Wolitzky-Taylor et al., 2012). Two other studies, also conducted by the same research team, compared Internet-based CBT to live group CBT

in two papers (El Alaoui et al., 2013; Hedman et al., 2012). Even these two sets of studies, however, treated different anxiety disorder populations (mixed and social anxiety disorder for CBT vs. ACT; panic disorder and social anxiety disorder for ICBT vs. CBGT). Although we presume that many of these putative moderators may work similarly across the anxiety disorders, these discrepancies may be explained by possible disorder-specific differences that should be the focus of future research. In other words, large studies with mixed anxiety disorders could also examine whether these moderators have different effects dependent on anxiety disorder type (e.g., specific anxiety disorder by anxiety sensitivity by group interaction). Clearly, much larger samples would be needed to achieve this goal.

Another challenge is that many putative moderators were defined differently across studies. For example, whereas some studies explored whether comorbid disorders in general moderated outcomes, others explored comorbid *anxiety* disorders, and still others explored general anxiety *levels* dimensionally. Similarly, some looked at the effect of a (categorically defined) comorbid mood disorder, whereas others looked at dimensional depressive symptoms. The measurement of anxiety disorder severity, in particular, was operationalized in several different ways. Some papers examined *overall* severity, some examined *specific symptom* severities, others measured severity in terms of quality of life, and still others assessed severity at the level of behavioral functioning. Even when moderators were defined similarly, they were often measured in different ways. For example, results were different when experiential avoidance was measured using self-report and using a laboratory task in the same study. The diverse operationalizations and measurement of moderating variables may have hindered our ability to detect consistent findings across studies.

In addition, many of the variants of CBT that were directly compared to each other, including those therapies we considered distinct from (traditional) CBT in this review such as ACT or MBSR, are not that distinct from each other (Arch & Craske, 2008; Mennin, Ellard, Fresco, & Gross, 2013). Therefore it is feasible that the same type of person who responds well to one treatment would also respond well to the comparative treatment. Similarly, all CBT therapies are not necessarily the same; CBT is a broad term encompassing many different packages of interventions that fall under the same umbrella. For example, studies that employ traditional CBT are in fact employing diverse multi-component treatment packages that could be rather distinct in content, focus, and delivery, that is, non-equivalent. Identifying consistent moderators for a particular treatment is thus challenged by the fact that treatment is a moving target.

Another possibility is that treatment moderators are particularly sensitive to the idiosyncrasies of the treatment sample, challenging replication efforts. However, the modest number, methodological quality, and statistical power of existing moderator studies suggests that the state of the research is too nascent to draw such conclusions.

As we move forward, it will be important to consider how the lack of consistency has affected the validity of findings to date, and how to improve validity and reliability in future studies. Thus far, no two studies have explored both the same treatments and the same populations. Even if this were the case, due to the diversity in definitions and measurement of moderators, we cannot assume that a moderator in one study is equivalent to the same moderator in another study. Likewise, we cannot assume that similar treatment conditions across studies are comparable, or that different treatment conditions are not. Despite these issues with validity and reliability, the importance of uncovering treatment moderators to develop personalized treatment for anxiety disorders, along with the consistent moderator findings observed in limited areas, supports our view that such efforts are worthwhile.

4.3. Methodological limitations

In addition to the aforementioned challenges, the anxiety disorder moderator literature is often beset by statistical and sample size limitations that compromise hypothesis testing, interpretation, and replication efforts. The statistical and methodological limitations of the anxiety disorder treatment moderator literature are worth noting, because they inform this review and the future directions proposed next. First, a striking number of statistical moderation tests were extremely difficult to decipher due to minimally reported or missing statistical results. As the superscripts in Table 1 note, even those papers that report statistics for moderators often did so in an incomplete or inconsistent manner from one putative moderator to the next, or describe findings without providing the statistics to back them up.

Second and relatedly, as evident in Fig. 3, the methodological quality and sample sizes of the 24 reviewed studies varied widely. If sample size is also considered, only one study garnered both the highest statistical rating and had relatively robust samples sizes. Nearly one-third (7/24) of studies did not follow up significant moderator by treatment condition interaction findings with any post-hoc tests that explored the nature of the interaction. Some examined interactions within subgroups of the moderator rather than within the entire sample. Although this is a statistically adequate way to examine the nature of the interaction, it is often significantly underpowered due to the small sample sizes within each subgroup selected for a post-hoc test. Because the moderator analyses generally represented secondary or exploratory analyses, they were often too underpowered to detect anything beyond a large or very large effect. Underpowered analyses become still more underpowered when subgroups are used for post hoc testing. In sum, treatment moderation tests were dramatically underpowered and rarely reflected the recommended large sample sizes often required to detect moderation (see McClelland & Judd, 1993). Despite the fact that *p*-values are affected by sample size, the majority of studies relied solely on *p*-values and failed to report effect sizes, which would also negatively impact their ability to detect smaller effects.

Detection of effects was also likely limited by the relatively homogenous samples used, which minimized the total amount of variability. For example, many studies excluded participants with severe depression or comorbidities such as bipolar disorder or substance use disorders, thus limiting the conclusions that could be drawn about therapy performance across the full range of patients. Few to no patients scoring at the upper limits on some putative moderators likely limited statistical power for testing moderation (see McClelland & Judd, 1993). Beyond basic methodological concerns, as noted, very few tested for non-linear relationships, likely limiting the detection of effects.

Finally, no studies tested moderators a priori by assigning participants to groups after assessing the moderator of interest, thereby precluding conclusions about causal effects or about the clinical significance of a statistically significant result.
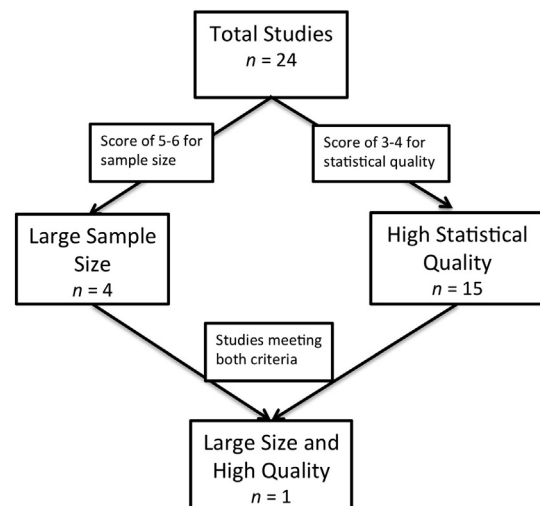

Fig. 3. Statistical quality and power of extant moderator studies.

In addition, we note limitations of the current review. The nascent state of the anxiety disorder treatment moderation literature precluded a quantified meta-analysis (see *Methods*); we thus were limited to a qualitative content review and quantitative methodological ratings. We had originally hoped to synthesize the literature and provide direct recommendations for clinical practice, but the state of the literature remains too nascent and inconsistent to realize this more ambitious aim. Finally, we were limited to articles written in English, and although we conducted a systematic search in multiple databases, it remains possible that we overlooked studies.

4.4. Recommendations for future directions

The goal of this line of research is to be able to inform clinicians about how to match patients to treatments that will be most effective for each person. To achieve this aim, the methodological quality and consistency of moderator research must improve. Gleaning lessons from our review of the extant literature, we thus propose recommendations for future moderator studies to help advance this important research. Specifically, we recommend that:

1. All studies comparing two different treatments for anxiety disorders should conduct at least basic moderation analyses if sufficient power exists to detect at least medium to large moderator effects. Most anxiety disorder treatment studies to date have investigated only treatment outcomes or at most, treatment predictors. Prior to examining treatment predictors, studies should *first examine treatment moderators*. Only if moderator findings are null should baseline variables then be tested as predictors. A habit of examining moderators first, prior to (or at least concurrently with) examining predictors and reporting all findings (both null and significant) will dramatically increase the number of papers that inform the personalized treatment of anxiety disorders.

   In addition to conducting moderator analyses on any sufficiently powered study that compares two or more treatments, designing studies whose primary aim is to investigate or replicate treatment moderators is essential to the development of this field. Studies designed with tests of moderation as the primary aim would avoid many of the methodological problems and limitations that are ubiquitous in the extant literature.

2. Putative moderators should be tested for quadratic (non-linear) relationships with outcomes in addition to linear relationships (see Wolitzky-Taylor et al., 2012). If quadratic effects are non-significant, linear effects alone should be assessed. The few studies that examined quadratic moderator by condition interactions reported that multiple moderator findings were characterized by a quadratic effect (Arch & Ayers, 2013; Craske et al., 2014; Wolitzky-Taylor et al., 2012). The implicit assumption that baseline variables relate to outcomes in a strictly linear manner is one possible reason why significant moderator findings and replication efforts continue to elude researchers.

3. Moderator statistical analyses, both null and significant findings, should be explained and reported as visibly and in as much detail as possible. In many papers, we found it difficult to ascertain the nature of the analyses that were conducted and whether they met the criteria for treatment moderation. Additionally, at times putative moderators were presented in the methods or discussion section but were not reported in the results section, rendering unclear whether they were tested as moderators but were non-significant, or whether they were tested as moderators at all. Consistent null findings, particularly from well-powered studies, have important clinical implications and also demonstrate honesty in research reporting. In sum, all moderator findings should be clearly and carefully reported.

   In addition, both effect sizes and *p*-values should be reported. Particularly in the case of underpowered studies, interpreting *p*-values alone is often insufficient to detect smaller, but potentially important, effects. Multiple independent findings of non-significant *p*-values with positive effect sizes could point towards a meaningful, small to moderate effect that would otherwise have been overlooked.

4. When choosing which moderators to test, we recommend first considering what has already been tested. Including valid baseline measures that have been examined within previous moderator studies strongly facilitates replication efforts. Standardizing and clarifying the operational definitions of the most commonly tested moderators (e.g., testing "DSM-5 comorbid mood disorders defined as major depressive disorder and dysthymia" rather than "comorbidity" as a moderator) would also facilitate replication efforts and enhance consistency across studies.

5. To limit Type I error, following Kraemer et al. (2006) we suggest that studies should limit moderator tests to sociodemographic variables and a limited number of a priori selected clinical characteristics of interest. Several additional points regarding moderator selection are worth considering. First, it can be helpful to note whether putative moderators are associated with underlying theories of the therapies being tested or are atheoretical in nature. Both theory-driven (e.g., degree of maladaptive beliefs in traditional CBT) and atheoretical (education, sex) moderators can be appropriate. Noting a priori moderator hypotheses clearly (for theory-driven moderators) or the lack thereof (for exploratory and atheoretical moderators) helps to guide the interpretation of moderator findings. Although both theory-driven and atheoretical moderators are important to explore, testing theory-driven moderators, particularly when comparing two treatments hypothesized to work through different mechanisms, is particularly informative. Second, to inform the current emphasis on transdiagnostic approaches and NIMH's Research Domain Criteria (RDoC), we also recommend selecting putative moderators that are applicable across multiple disorders (e.g., comorbid mood disorders) as an equal priority to disorder-specific ones (e.g., beliefs regarding the physiological consequences of panic attacks). Broader moderators may translate more readily into clinical practice. Third, physiological or biological moderation tests were scant and represent an important area for future inquiry. This area would benefit from theoretical and empirical work that guides predictions about which biological factors could be expected to moderate anxiety disorder treatment outcomes, why, and how.

6. Future researchers should consider using recently proposed statistical approaches that compare and combine multiple moderators in a manner that better facilitates treatment matching, such as Kraemer's (2013) recently proposed approach to developing composite moderators. In addition, after identifying specific treatment moderators based on methods proposed here, future studies should more definitively assess the clinical utility of such moderators. This could be achieved by assessing moderators at baseline and subsequently assigning patients to treatment based on the results of this assessment. Assigning participants to treatments based on moderator assessment could be used in at least two important ways: a) to ensure similar distributions of the moderator in each group; and b) to test whether assigning clients to a specific treatment based on their level of the moderator makes a difference in clinical practice.

7. Finally, this represents a list of recommendations for future studies, rather than as a methodological "how-to" guide for conducting moderator analyses (see Kraemer et al., 2002, for an excellent guide to conducting moderator analyses). Nonetheless, given the observed confusion on this point, we provide a brief outline of the statistical steps for conducting a basic treatment moderator analysis:

   a. The first step is to conduct a moderator by condition interaction test (see Baron & Kenny, 1986; Kraemer et al., 2002). We identified

multiple papers purporting to investigate treatment moderation that investigated purported moderators only *within* each treatment condition and never compared *across* treatment conditions (e.g., Brody et al., 1998; Cottraux et al., 1990; Hoexter et al., 2013; Nordgreen et al., 2012). Limiting analyses to a single treatment condition is sufficient to conclude within-group prediction, but not treatment moderation.

b. Following the initial interaction test, the next step is to conduct between-group (often post hoc) tests of simple effects at different points along the moderator (e.g., at the mean, at 1 *SD* above and below the mean for a continuous moderator or at each level of a categorical moderator; see Aiken & West, 1991 and Holmbeck, 2002). This allows researchers to explore the nature of the interaction. This analysis is best conducted, whenever possible, using methods that allow for use of all of the data (such as hierarchical linear modeling in instances of three or more assessment points), thereby maximizing statistical power. This step is essential to elucidate if and how the interaction is meaningful for a given treatment comparison.

c. Although within-group post hoc tests of simple effects are not necessary to assess moderation, which is at essence a between-groups assessment, we recommend them because they can provide useful prognostic information.

5. Conclusion

Although there has been growing interest in the identification of treatment moderators, relatively few studies have examined moderators of CBT treatment for adult anxiety disorders. The majority of these studies have methodological or sample size limitations, demonstrating the need for future methodological improvements and larger sample sizes. Despite this, some findings (albeit in need of replication) are beginning to emerge. To continue working towards the goal of personalized treatment of anxiety disorders, we must prioritize the identification and characterization of treatment moderators using the best methods available.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.cpr.2015.02.004.

Conflict of interest

We wish to confirm that there are no known conflicts of interest associated with this publication.

Role of funding sources

There has been no significant financial support for this work that could have influenced its outcome.

Contributors

We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us.

We confirm that we have given due consideration to the protection of intellectual property associated with this work and that there are no impediments to publication, including the timing of publication, with respect to intellectual property. In so doing we confirm that we have followed the regulations of our institutions concerning intellectual property.

We understand that the Corresponding Author is the sole contact for the Editorial process (including Editorial Manager and direct communications with the office). She is responsible for communicating with the other authors about progress, submissions of revisions and final approval of proofs. We confirm that we have provided a current, correct email address which is accessible by the Corresponding Author and which has been configured to accept email from CPR@elsevier.com.

Signed by all authors as follows:
Joanna J. Arch 2/06/15
Rebecca L. Schneider 2/06/15
Kate B. Wolitzky-Taylor 2/06/15

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions.* Newbury Park, Calif: Sage Publications.

Andersson, G., Carlbring, P., & Grimlund, A. (2008). Predicting treatment outcome in internet versus face to face treatment of panic disorder. *Computers in Human Behavior, 24*(5), 1790–1801. http://dx.doi.org/10.1016/j.chb.2008.02.003.

Arch, J. J., & Ayers, C. R. (2013). Which treatment worked better for whom? Moderators of group cognitive behavioral therapy versus adapted mindfulness based stress reduction for anxiety disorders. *Behaviour Research and Therapy, 51*(8), 434–442. http://dx.doi.org/10.1016/j.brat.2013.04.004.

Arch, J. J., & Craske, M. G. (2008). Acceptance and commitment therapy and cognitive behavioral therapy for anxiety disorders: Different treatments, similar mechanisms? *Clinical Psychology: Science and Practice, 15*(4), 263–279.

Arch, J. J., & Craske, M. G. (2009). First-line treatment: A critical appraisal of cognitive behavioral therapy developments and alternatives. *Psychiatric Clinics of North America, 32*(3), 525–547. http://dx.doi.org/10.1016/j.psc.2009.05.001.

Barlow, D. H., Allen, L. B., & Choate, M. L. (2004). Toward a unified treatment for emotional disorders. *Behavior Therapy, 35*(2), 205–230. http://dx.doi.org/10.1016/s0005-7894(04)80036-4.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173. http://dx.doi.org/10.1037/0022-3514.51.6.1173.

Brody, A. L., Saxena, S., Schwartz, J. M., Stoessel, P. W., Maidment, K., Phelps, M. E., et al. (1998). FDG-PET predictors of response to behavioral therapy and pharmacotherapy in obsessive compulsive disorder. *Psychiatry Research, 84*(1), 1–6. http://dx.doi.org/10.1016/s0925-4927(98)00041-9.

Burton, M., Schmertz, S. K., Price, M., Masuda, A., & Anderson, P. L. (2013). The relation between mindfulness and fear of negative evaluation over the course of cognitive behavioral therapy for social anxiety disorder: The relation between mindfulness and fear. *Journal of Clinical Psychology, 69*(3), 222–228. http://dx.doi.org/10.1002/jclp.21929.

Butler, A., Chapman, J., Forman, E., & Beck, A. (2006). The empirical status of cognitive-behavioral therapy: A review of meta-analyses. *Clinical Psychology Review, 26*(1), 17–31. http://dx.doi.org/10.1016/j.cpr.2005.07.003.

Chambless, D. L., Baker, M. J., Baucom, D. H., Beutler, L. E., Calhoun, K. S., Crits-Christoph, P., et al. (1998). Update on empirically validated therapies, II. *Clinical Psychologist, 51*(1), 3–16. http://dx.doi.org/10.1037/e619622010-001.

Chambless, D. L., & Ollendick, T. H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology, 52*, 685–716. http://dx.doi.org/10.1146/annurev.psych.52.1.685.

Chambless, D. L., Sanderson, W. C., Shoham, V., Bennett Johnson, S., Pope, K. S., Crits-Christoph, P., et al. (1996). An update on empirically validated therapies. *The Clinical Psychologist, 49*, 5–18. http://dx.doi.org/10.1037/e555332011-003.

Cottraux, J., Mollard, E., Bouvard, M., Marks, I., Sluys, M., Nury, A. M., et al. (1990). A controlled study of fluvoxamine and exposure in obsessive–compulsive disorder. *International Clinical Psychopharmacology, 5*(1), 17–30. http://dx.doi.org/10.1097/00004850-199001000-00002.

Craske, M. G., Niles, A. N., Burklund, L. J., Wolitzky-Taylor, K. B., Vilardaga, J. C. P., Arch, J. J., et al. (2014). Randomized controlled trial of cognitive behavioral therapy and acceptance and commitment therapy for social phobia: Outcomes and moderators. *Journal of Consulting and Clinical Psychology, 82*(6), 1034–1048. http://dx.doi.org/10.1037/a0037212.

Craske, M. G., Rauch, S. L., Ursano, R., Prenoveau, J., Pine, D. S., & Zinbarg, R. E. (2009). What is an anxiety disorder? *Depression and Anxiety, 26*(12), 1066–1085. http://dx.doi.org/10.1002/da.20633.

D'Alcante, C. C., Diniz, J. B., Fossaluza, V., Batistuzzo, M. C., Lopes, A. C., Shavitt, R. G., et al. (2012). Neuropsychological predictors of response to randomized treatment in obsessive–compulsive disorder. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 39*(2), 310–317. http://dx.doi.org/10.1016/j.pnpbp.2012.07.002.

Davies, C. D., Niles, A. N., Pittig, A., Arch, J. J., & Craske, M. G. (2015). Physiological and behavioral indices of emotion dysregulation as predictors of outcome from cognitive behavioral therapy and acceptance and commitment therapy for anxiety. *Journal of Behavior Therapy and Experimental Psychiatry, 46*, 35–43. http://dx.doi.org/10.1016/j.jbtep.2014.08.002.

de Haan, E., van Oppen, P., Balkom, A., Spinhoven, P., Hoogduin, K. A. L., & van Dyck, R. (1997). Prediction of outcome and early vs. late improvement in OCD patients treated with cognitive behaviour therapy and pharmacotherapy. *Acta Psychiatrica Scandinavica, 96*(5), 354–361. http://dx.doi.org/10.1111/j.1600-0447.1997.tb09929.x.

De Kleine, R. A., Hendriks, G. -J., Kusters, W. J. C., Broekman, T. G., & van Minnen, A. (2012). A randomized placebo-controlled trial of D-cycloserine to enhance exposure therapy for posttraumatic stress disorder. *Biological Psychiatry, 71*(11), 962–968. http://dx.doi.org/10.1016/j.biopsych.2012.02.033.

De Kleine, R. A., Hendriks, G. -J., Smits, J. A. J., Broekman, T. G., & van Minnen, A. (2014). Prescriptive variables for ᴅ-cycloserine augmentation of exposure therapy for post-traumatic stress disorder. *Journal of Psychiatric Research*, 48(1), 40–46. http://dx.doi.org/10.1016/j.jpsychires.2013.10.008.

Dow, M. G. T., Kenardy, J. A., Johnston, D. W., Newman, M. G., Taylor, C. B., & Thomson, A. (2007). Prognostic indices with brief and standard CBT for panic disorder: II. Moderators of outcome. *Psychological Medicine*, 37(10). http://dx.doi.org/10.1017/S0033291707000682.

El Alaoui, S., Hedman, E., Ljótsson, B., Bergström, J., Andersson, E., & Rück, C. (2013). Predictors and Moderators of Internet-and Group-Based Cognitive Behaviour Therapy for Panic Disorder. *PloS one*, 8(11), e79024.

Eskildsen, A., Hougaard, E., & Rosenberg, N. K. (2010). Pre-treatment patient variables as predictors of drop-out and treatment outcome in cognitive behavioural therapy for social phobia: A systematic review. *Nordic Journal of Psychiatry*, 64(2), 94–105. http://dx.doi.org/10.3109/08039480903426929.

Ginsburg, G. S., & Willard, H. F. (2009). Genomic and personalized medicine: Foundations and applications. *Translational Research*, 154(6), 277–287. http://dx.doi.org/10.1016/j.trsl.2009.09.005.

Hansen, B., Vogel, P. A., Stiles, T. C., & Gunnar Götestam, K. (2007). Influence of co-morbid generalized anxiety disorder, panic disorder and personality disorders on the outcome of cognitive behavioural treatment of obsessive–compulsive disorder. *Cognitive Behaviour Therapy*, 36(3), 145–155. http://dx.doi.org/10.1080/16506070701259374.

Hedman, E., Andersson, E., Ljótsson, B., Andersson, G., Andersson, E., Schalling, M., et al. (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. http://dx.doi.org/10.1111/j.1600-0447.2012.01834.x.

Hicks, T. V., Leitenberg, H., Barlow, D. H., Gorman, J. M., Shear, M. K., & Woods, S. W. (2005). Physical, mental, and social catastrophic cognitions as prognostic factors in cognitive-behavioral and pharmacological treatments for panic disorder. *Journal of Consulting and Clinical Psychology*, 73(3), 506–514. http://dx.doi.org/10.1037/0022-006X.73.3.506.

Hoexter, M. Q., Dougherty, D. D., Shavitt, R. G., D'Alcante, C. C., Duran, F. L. S., Lopes, A. C., et al. (2013). Differential prefrontal gray matter correlates of treatment response to fluoxetine or cognitive-behavioral therapy in obsessive–compulsive disorder. *European Neuropsychopharmacology*, 23(7), 569–580. http://dx.doi.org/10.1016/j.euroneuro.2012.06.014.

Hofmann, S. G., Asnaani, A., Vonk, I. J. J., Sawyer, A. T., & Fang, A. (2012). The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive Therapy and Research*, 36(5), 427–440. http://dx.doi.org/10.1007/s10608-012-9476-1.

Hofmann, S. G., Sawyer, A. T., Witt, A. A., & Oh, D. (2010). The effect of mindfulness-based therapy on anxiety and depression: A meta-analytic review. *Journal of Consulting and Clinical Psychology*, 78(2), 169–183. http://dx.doi.org/10.1037/a0018555.

Hofmann, S. G., & Smits, J. A. (2008). Cognitive-behavioral therapy for adult anxiety disorders: A meta-analysis of randomized placebo-controlled trials. *The Journal of Clinical Psychiatry*, 69(4), 621. http://dx.doi.org/10.4088/jcp.v69n0415.

Hohagen, F., Winkelmann, G., Rasche-Rüchle, H., Hand, I., König, A., Münchau, N., et al. (1998). Combination of behaviour therapy with fluvoxamine in comparison with behaviour therapy and placebo. Results of a multicentre study. *The British Journal of Psychiatry. Supplement*(35), 71–78.

Holmbeck, G. N. (2002). Post-hoc probing of significant moderational and mediational effects in studies of pediatric populations. *Journal of Pediatric Psychology*, 27(1), 87–96. http://dx.doi.org/10.1093/jpepsy/27.1.87.

Insel, T. R. (2009). Translating scientific opportunity into public health impact: A strategic plan for research on mental illness. *Archives of General Psychiatry*, 66(2), 128–133. http://dx.doi.org/10.1001/archgenpsychiatry.2008.540.

Joutsenniemi, K., Laaksonen, M. A., Knekt, P., Haaramo, P., & Lindfors, O. (2012). Prediction of the outcome of short- and long-term psychotherapy based on socio-demographic factors. *Journal of Affective Disorders*, 141(2–3), 331–342. http://dx.doi.org/10.1016/j.jad.2012.03.027.

Keeley, M. L., Storch, E. A., Merlo, L. J., & Geffken, G. R. (2008). Clinical predictors of response to cognitive-behavioral therapy for obsessive–compulsive disorder. *Clinical Psychology Review*, 28(1), 118–130. http://dx.doi.org/10.1016/j.cpr.2007.04.003.

Klass, E. T., Milrod, B. L., Leon, A. C., Kay, S. J., Schwalberg, M., Li, C., et al. (2009). Does interpersonal loss preceding panic disorder onset moderate response to psychotherapy? An exploratory study. *The Journal of Clinical Psychiatry*, 70(3), 406–411. http://dx.doi.org/10.4088/JCP.08m04248.

Knopp, J., Knowles, S., Bee, P., Lovell, K., & Bower, P. (2013). A systematic review of predictors and moderators of response to psychological therapies in OCD: Do we have enough empirical evidence to target treatment? *Clinical Psychology Review*, 33(8), 1067–1081. http://dx.doi.org/10.1016/j.cpr.2013.08.008.

Kraemer, H. C. (2013). Discovering, comparing, and combining moderators of treatment on outcome after randomized clinical trials: A parametric approach. *Statistics in Medicine*, 32, 1964–1973. http://dx.doi.org/10.1002/sim.5734.

Kraemer, H. C., Frank, E., & Kupfer, D. J. (2006). Moderators of treatment outcomes: Clinical, research, and policy importance. *JAMA*, 296(10), 1286–1289. http://dx.doi.org/10.1001/jama.296.10.1286.

Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry*, 59(10), 877. http://dx.doi.org/10.1001/archpsyc.59.10.877.

Luborsky, L., Auerbach, A. H., Chandler, M., Cohen, J., & Bachrach, H. M. (1971). Factors influencing the outcome of psychotherapy: A review of quantitative research. *Psychological Bulletin*, 75(3), 145.

Maher, M. J., Huppert, J. D., Chen, H., Duan, N., Foa, E. B., Liebowitz, M. R., et al. (2010). Moderators and predictors of response to cognitive-behavioral therapy

augmentation of pharmacotherapy in obsessive–compulsive disorder. *Psychological Medicine*, 40(12), 2013–2023. http://dx.doi.org/10.1017/s0033291710000620.

McClelland, G. H., & Judd, C. M. (1993). Statistical difficulties of detecting interactions and moderator effects. *Psychological Bulletin*, 114(2), 376–390. http://dx.doi.org/10.1037/0033-2909.114.2.376.

Mennin, D. S., Ellard, K. K., Fresco, D. M., & Gross, J. J. (2013). United we stand: Emphasizing commonalities across cognitive-behavioral therapies. *Behavior Therapy*, 44(2), 234–248. http://dx.doi.org/10.1016/j.beth.2013.02.004.

Mennin, D. S., & Heimberg, R. G. (2000). The impact of comorbid mood and personality disorders in the cognitive-behavioral treatment of panic disorder. *Clinical Psychology Review*, 20(3), 339–357. http://dx.doi.org/10.1016/s0272-7358(98)00095-6.

Meuret, A. E., Hofmann, S. G., & Rosenfield, D. (2010). Catastrophic appraisal and perceived control as moderators of treatment response in panic disorder. *International Journal of Cognitive Therapy*, 3(3), 262–277. http://dx.doi.org/10.1521/ijct.2010.3.3.262.

Milrod, B. L., Leon, A. C., Barber, J. P., Markowitz, J. C., & Graf, E. (2007). Do comorbid personality disorders moderate panic-focused psychotherapy? An exploratory examination of the American Psychiatric Association practice guideline. *The Journal of Clinical Psychiatry*, 68(6), 885–891. http://dx.doi.org/10.4088/jcp.v68n0610.

Mululo, S. C. C., de Menezes, G. B., Vigne, P., & Fontenelle, L. F. (2012). A review on predictors of treatment outcome in social anxiety disorder. *Revista Brasileira de Psiquiatria (São Paulo, Brazil: 1999)*, 34(1), 92–100. http://dx.doi.org/10.1016/s1516-4446(12)70016-8.

Newman, M. G., & Fisher, A. J. (2013). Mediated moderation in combined cognitive behavioral therapy versus component treatments for generalized anxiety disorder. *Journal of Consulting and Clinical Psychology*, 81(3), 405–414. http://dx.doi.org/10.1037/a0031690.

Niles, A. N., Mesri, B., Burklund, L. J., Lieberman, M. D., & Craske, M. G. (2013). Attentional bias and emotional reactivity as predictors and moderators of behavioral treatment for social phobia. *Behaviour Research and Therapy*, 51(10), 669–679. http://dx.doi.org/10.1016/j.brat.2013.06.005.

Nilsen, T. S., Eisemann, M., & Kvernmo, S. (2013). Predictors and moderators of outcome in child and adolescent anxiety and depression: A systematic review of psychological treatment studies. *European Child & Adolescent Psychiatry*, 22(2), 69–87. http://dx.doi.org/10.1007/s00787-012-0316-3.

Nordgreen, T., Havik, O. E., Öst, L. G., Furmark, T., Carlbring, P., & Andersson, G. (2012). Outcome predictors in guided and unguided self-help for social anxiety disorder. *Behaviour Research and Therapy*, 50(1), 13–21. http://dx.doi.org/10.1016/j.brat.2011.10.009.

Norton, P. J., & Price, E. C. (2007). A meta-analytic review of adult cognitive-behavioral treatment outcome across the anxiety disorders. *The Journal of Nervous and Mental Disease*, 195(6), 521–531. http://dx.doi.org/10.1097/01.nmd.0000253843.70149.9a.

Olatunji, B. O., Cisler, J. M., & Deacon, B. J. (2010). Efficacy of cognitive behavioral therapy for anxiety disorders: A review of meta-analytic findings. *Psychiatric Clinics of North America*, 33(3), 557–577. http://dx.doi.org/10.1016/j.psc.2010.04.002.

Olatunji, B. O., Davis, M. L., Powers, M. B., & Smits, J. A. J. (2013). Cognitive-behavioral therapy for obsessive–compulsive disorder: A meta-analysis of treatment outcome and moderators. *Journal of Psychiatric Research*, 47(1), 33–41. http://dx.doi.org/10.1016/j.jpsychires.2012.08.020.

Pampaloni, I., Bruscoli, M., & Pallanti, S. (2004). Obsessive–compulsive disorder: Clinical response predictors. *Clinical Neuropsychiatry: Journal of Treatment Evaluation, 1*(1), 52–58.

Pan, D., Huey, S. J., & Hernandez, D. (2011). Culturally adapted versus standard exposure treatment for phobic Asian Americans: Treatment efficacy, moderators, and predictors. *Cultural Diversity and Ethnic Minority Psychology*, 17(1), 11–22. http://dx.doi.org/10.1037/a0022534.

Rodebaugh, T. L., Heimberg, R. G., Schultz, L. T., & Blackmore, M. (2010). The moderated effects of video feedback for social anxiety disorder. *Journal of Anxiety Disorders*, 24(7), 663–671. http://dx.doi.org/10.1016/j.janxdis.2010.04.007.

Simon, G. E., & Perlis, R. H. (2010). Personalized medicine for depression: Can we match patients with treatments? *American Journal of Psychiatry*, 167(12), 1445–1455. http://dx.doi.org/10.1176/appi.ajp.2010.09111680.

Smits, J. A. J., Hofmann, S. G., Rosenfield, D., DeBoer, L. B., Costa, P. T., Simon, N. M., et al. (2013). ᴅ-cycloserine augmentation of cognitive behavioral group therapy of social anxiety disorder: Prognostic and prescriptive variables. *Journal of Consulting and Clinical Psychology*. http://dx.doi.org/10.1037/a0034120.

Solvason, H. B., Ernst, H., & Roth, W. (2003). Predictors of response in anxiety disorders. *Psychiatric Clinics of North America*, 26(2), 411–433. http://dx.doi.org/10.1016/S0193-953X(03)00027-3.

Steketee, G., & Shapiro, L. J. (1995). Predicting behavioral treatment outcome for agoraphobia and obsessive compulsive disorder. *Clinical Psychology Review*, 15(4), 317–346. http://dx.doi.org/10.1016/0272-7358(95)00017-j.

Taylor, S. (2003). Outcome predictors for three PTSD treatments: Exposure therapy, EMDR, and relaxation training. *Journal of Cognitive Psychotherapy*, 17(2), 149–162. http://dx.doi.org/10.1891/088983903780906219.

Taylor, S., Abramowitz, J. S., & McKay, D. (2012). Non-adherence and non-response in the treatment of anxiety disorders. *Journal of Anxiety Disorders*, 26(5), 583–589. http://dx.doi.org/10.1016/j.janxdis.2012.02.010.

Thrasher, S., Power, M., Morant, N., Marks, I., & Dalgleish, T. (2010). Social support moderates outcome in a randomized controlled trial of exposure therapy and (or) cognitive restructuring for chronic posttraumatic stress disorder. *Canadian Journal of Psychiatry. Revue Canadienne de Psychiatrie*, 55(3), 187–190.

Tolin, D. F. (2010). Is cognitive-behavioral therapy more effective than other therapies? A meta-analytic review. *Clinical Psychology Review*, 30(6), 710–720. http://dx.doi.org/10.1016/j.cpr.2010.05.003.

Tyrer, P., Seivewright, N., Ferguson, B., Murphy, S., & Johnson, A. L. (1993). The Nottingham study of neurotic disorder. Effect of personality status on response to drug treatment, cognitive therapy and self-help over two years. *The British Journal of Psychiatry, 162*(2), 219–226. http://dx.doi.org/10.1192/bjp.162.2.219.

Walters, K., Rait, G., Griffin, M., Buszewicz, M., & Nazareth, I. (2012). Recent trends in the incidence of anxiety diagnoses and symptoms in primary care. *PLoS ONE, 7*(8), e41670. http://dx.doi.org/10.1371/journal.pone.0041670.

Westra, H. A., Arkowitz, H., & Dozois, D. J. A. (2009). Adding a motivational interviewing pretreatment to cognitive behavioral therapy for generalized anxiety disorder: A preliminary randomized controlled trial. *Journal of Anxiety Disorders, 23*(8), 1106–1117. http://dx.doi.org/10.1016/j.janxdis.2009.07.014.

Wolitzky-Taylor, K. B., Arch, J. J., Rosenfield, D., & Craske, M. G. (2012). Moderators and non-specific predictors of treatment outcome for anxiety disorders: A comparison of cognitive behavioral therapy to acceptance and commitment therapy. *Journal of Consulting and Clinical Psychology, 80*(5), 786–799. http://dx.doi.org/10.1037/a0029418.

Zettle, R. D. (2003). Acceptance and commitment therapy (ACT) vs. systematic desensitization in treatment of mathematics anxiety. *The Psychological Record, 53*(2), 197–215.