RUNNING HEAD: Meta-analysis of the efficacy of behavioral activation

Looking beyond depression: A meta-analysis of the effect of behavioral activation on depression, anxiety, and activation

Aliza T Stein[1], Emily Carl[1], Pim Cuijpers[2], Eirini Karyotaki[2,3], & Jasper A.J. Smits[1]

[1]University of Texas at Austin

[2]Vrije Universiteit Amsterdam

[3]Harvard Medical School

Address correspondence to:
Aliza Stein, Department of Psychology and Institute for Mental Health Research, University of Texas at Austin, 305 E. 23rd St., E9000, Austin, TX 78712. Email: atstein@utexas.edu

Financial support: This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Conflicts of Interest: Authors Aliza T Stein, Emily Carl, Pim Cuijpers, and Eirini Karyotaki declare that they have no conflicts of interest. Jasper A.J. Smits has received monetary compensation for his work as a clinical advisor to Big Health, Ltd.

Word count (text only, excluding abstract, references, tables/figures): 4,499

Abstract

Background: Depression is a prevalent and impairing condition. Behavioral Activation (BA) is a parsimonious, cost-effective, and easily disseminated psychological intervention for depression. The present meta-analysis expanded on existing literature supporting the efficacy of BA for depression by examining the effects of BA on additional relevant outcomes for patients with depression, namely the reduction in anxiety symptoms and increase in activation.

Methods: Randomized controlled trials of BA for depression compared to active and inactive control were identified via systematic review. Effect sizes using Hedges's *g* were calculated for each outcome compared to both active and inactive control using random effects models. Subgroup analyses were used to examine the inclusion of a discussion of values as a moderator of depression symptom outcome in BA.

Results: Twenty-eight studies were included. Meta-analyses of symptom change between groups from baseline-to-post intervention indicated that BA outperformed inactive control for depression (g = 0.83), anxiety (g = 0.37), and activation (g = 0.64). Effect sizes did not differ significantly from active control for depression (g = 0.15), anxiety (g = 0.03), activation (g = 0.04), or for subgroup analyses. Study quality was generally low, and evidence of publication bias was present.

Conclusions: Results support the efficacy of BA relative to inactive control for reducing symptoms of depression and anxiety in addition to increasing activation. BA did not differ from active control on these outcomes. These findings support the use of BA for the treatment of depression, yet call for continued high quality evaluations of BA for depression.

Beyond improvement in depression: A meta-analysis of the efficacy of behavioral activation

Major depressive disorder (depression) is the most prevalent mental health disorder, with a lifetime prevalence rate of approximately 17% (Kessler *et al.* 2005). According to the World Health Organization, depression is a leading cause of disability worldwide, affecting more than 300 million people across the lifespan (World Health Organization, 2017). Although there are a variety of empirically supported treatments for depression, Behavioral Activation (BA) is a parsimonious, cost-effective, and easily disseminated psychotherapy for depression (Richards *et al.* 2016).

According to behavioral models of depression, depressive symptoms develop as a result of a reduction in experiences with positive outcomes (RCPR; response-contingent positive reinforcement) coupled with an increase in negative or aversive events (Ferster 1973). This reduction in RCPR is exacerbated by diminished reinforcement of positive outcomes, a consequence of depressed mood (MacPhillamy & Lewinsohn 1974; Lewinsohn *et al.* 1980), thereby creating a vicious cycle of decreased activity and low mood. Behavioral activation treatments have been developed as a means of breaking this negative feedback loop by increasing participation in rewarding activities (Martell *et al.* 2013). Through the process of increasing activity, the patient has more opportunities to come into contact with rewarding experiences, which is thought to decrease depression over time. Behavioral activation primarily involves self-monitoring and activity scheduling aimed at increasing overt behaviors that are intended to bring the individual into contact with positive reinforcers in the environment.

Previous meta-analyses have supported the efficacy of behavioral activation for reducing depressive symptoms in adults (Cuijpers *et al.* 2007; Mazzucchelli *et al.* 2009; Sturmey 2009; Ekers *et al.* 2014; Cuijpers *et al.* 2019b). For example, Ekers et al. (2014) showed that BA had

large antidepressant effects relative to control conditions and Barth et al. (2013) reported that changes in depression severity are not different in BA from those observed with other established interventions (e.g., interpersonal psychotherapy, cognitive behavior therapy, problem solving therapy).

Indeed, most individuals with depression experience co-occurring anxiety. In a large cohort study conducted in the Netherlands, an estimated 75% of individuals with depression had a lifetime comorbid anxiety disorder (Lamers et al., 2011). Accordingly, estimating the effects of BA on anxiety symptom severity has potential to guide clinicians in their treatment planning. Currently, there is limited information available to guide clinicians in selection psychotherapies for depression and co-occurring disorders. We included activation as an outcome in this meta-analysis because it is a core mechanistic target of BA and an estimate of target engagement has the potential to guide future efforts to optimize BA efficacy. We also tested whether the effect sizes on depression symptoms varied depending on the nature of the BA protocol. Specifically, more recent BA interventions have expanded the focus on incorporating a values assessment (Lejuez *et al.* 2001, 2011) in activity scheduling as a means to maximize activation and improve outcomes. Hence, we tested whether including a discussion of values in treatment resulted in better depression outcomes.

The present meta-analysis sought to expand on this literature by examining the effects of BA on additional relevant outcomes for patients with depression, namely the reduction in anxiety symptoms and the increase in activation. Specifically, the objective of this meta-analysis was to evaluate and summarize the current scientific knowledge regarding the effects of behavioral activation on (1) depression symptoms, (2) activation, and (3) anxiety symptoms compared to

both inactive and active control conditions. We further aimed to examine the moderating effect of incorporating a discussion of values into treatment on depression symptoms.

Method

Protocol and Registration

The PRISMA statement guidelines for systematic review were followed for this review (Moher *et al.* 2009). In accordance with these guidelines, the protocol for this review was pre-registered on PROSPERO (CRD42019124300).

Search Strategy

The Cochrane Library, PsycINFO, and PubMed databases were searched for relevant articles using the following search terms: ("behavioral activation" OR "behavioural activation" OR "activity scheduling" OR "pleasant events" OR "pleasant activities") AND depress*. Search parameters limited results to studies published before studies published in English before 01 February 2019 of peer-reviewed articles examining human subjects. We also used references extracted from previously published meta-analyses and systematic reviews and an existing database of psychotherapy studies of depression (for further details about this database refer to Cuijpers et al., 2019a).

Study Selection

Results from the search were uploaded into Covidence (Covidence systematic review software, 2019), which was used to manage data throughout the review process. Study titles and abstracts were first screened by two independent reviewers (AS and EC) for possible relevance to topic and eligibility criteria. Studies that were clearly not relevant or not meeting eligibility criteria based on title and abstract were excluded. Full text of the remaining studies was reviewed by two independent reviewers (AS and EC) and assessed for eligibility criteria. The two

reviewers discussed any discrepancies and final determinations were made through consensus. If consensus could not be reached, a third author assisted in determinations (JS). Data were then extracted by two independent researchers (AS and EC) into Covidence. The independent reviewers then compared the data extracted and resolved discrepancies as discussed above.

We employed the following inclusion criteria: (1) Studies were published in English in a peer-reviewed, scholarly journals before February 2019; (2) Studies aimed to evaluate the effect of behavioral activation on depression. For the purpose of this study, behavioral activation was defined as a time-limited treatment delivered individually (as opposed to group) by a trained clinician (as opposed to electronically), in which the primary treatment components were activity scheduling and self-monitoring for the purpose of reducing symptoms of depression; (3) Additional treatment components (e.g., cognitive restructuring, social skills training, problem solving) were acceptable as long as they were not a primary aim of the treatment (e.g., did not comprise entire treatment sessions); (4) Studies included a sample of adults (≥ 18 years of age) receiving treatment for depressed mood. We included studies that involved participants with medical and/or psychiatric comorbidity as long as the primary aim of the intervention was reduction in depression symptom severity. Hence, studies that tested the efficacy of combination interventions (e.g., targeting both depression and another co-occurring symptom or condition) were not included; (5) Employed a randomized controlled trial design (randomization must occur at the individual level), involving any type of control comparison condition. We did not include studies in which the control arm was another version of the same intervention (e.g., BA delivered individual vs group, internet vs in-person); (6) Administered psychometrically-sound measures of depression symptoms before and after the intervention.

Data Extraction

All data were extracted by two independent authors (AS and EC) and entered in Covidence (Covidence systematic review software, 2019). The data were then exported into a spreadsheet and checked independently by both authors prior to analyses. For all outcomes, sample sizes, means and standard deviations were extracted. When those data were not reported, authors were contacted with requests for additional information. If authors did not respond to two requests for the data, effect size data were extracted from either the original study report or a prior meta-analysis, when possible.

Quality Assessment

Two authors (AS and EC) independently rated risk of bias of all included studies using the Cochrane Collaboration's Risk of Bias tool (Higgins et al., 2011). Disagreements were resolved by discussion until consensus was reached. The following elements were rated: (1) Random sequence generation (selection bias), (2) Allocation concealment (selection bias), (3) Blinding of outcome assessment (detection bias), (4) Incomplete outcome data (attrition bias), (5) Selective reporting (reporting bias). Each element was rated to have a high, low, or unclear risk of bias. Because studies psychological interventions are typically not able to blind participants and personnel to condition, blinding of participants and personnel was not rated for this review.

Data synthesis and analysis

CMA version 3 (Borenstein, Hedges, Higgins, & Rothstein, 2013) was used to estimate controlled effect sizes using Hedges's *g* (Hedges & Olkin 2014). Hedges's *g* was used because it corrects for small sample sizes, providing a more accurate estimate (Hedges & Olkin, 2014). Hedges's *g* was calculated from extracted means and standard deviations and significance tests when appropriate. Hedges's *g* was interpreted using the same conventions as Cohen's *d,* with

0.2, 0.5, and 0.8 representing small, medium, and large effect sizes respectively. Pooled effect sizes were calculated for each outcome (Depression, Anxiety, and Activation) by control type (inactive or active) using random effects models. Random effects models were used due to the heterogeneity of included studies. Because we wanted to avoid including multiple comparison conditions from the same study in the effect size analyses (because this violates the independence assumption), we selected one control condition per analysis. If there were multiple inactive conditions, we opted to include the most stringent control condition (e.g., placebo/supportive counseling over waitlist). If there were multiple active control conditions, we favored interventions that were of the same treatment modality as BA (e.g., psychotherapy over pharmacotherapy). If there were multiple active conditions that were similar modalities to BA (e.g., multiple psychotherapy conditions), we chose the psychotherapy that was conceptually furthest from BA (e.g., psychodynamic over CBT). If there were multiple outcome measures for a single outcome, these measures were pooled for analyses.

Heterogeneity was assessed using the Q and I^2 statistics. Cochran's Q-test was used to test whether the observed variability between effect sizes is greater than what would be expected due to sampling error (Higgins *et al. ,* 2003). The I^2 statistic (Higgins *et al.*, 2003) is the percent of total variance explained by heterogeneity. Values can range from 0% (zero heterogeneity) to 100% (the difference in effect sizes is explained by sampling error), with 25%, 50%, and 75% indicating low, moderate, and high heterogeneity, respectively. I^2 was calculated using CMA and 95% confidence intervals were computed in excel using formulas provided by Borenstein et al (2011).

For each effect size analysis, a funnel plot was visually inspected for publication bias (Egger *et al.* 1997). Duval and Tweedie's trim and fill procedure was used to test for missing

studies (Duval & Tweedie 2000). Egger's test, which quantifies bias in the funnel plot, was also used to assess publication bias (Hedges & Olkin, 1985). It is important to note that these methods for testing heterogeneity and publication bias may be biased when there is a small number of studies and therefore should be interpreted cautiously. In an effort to mitigate this bias, multiple methods were used and confidence intervals were reported.

Moderation. Subgroup analyses were conducted to investigate the moderating effect of a discussion of values on depression symptom outcomes in BA versus inactive control. A mixed effect model was used, in which subgroups were pooled based on a presence or absence of a discussion on values for the random effects model. Differences between subgroups were tested using the fixed effects model.

Results

Search results and trial characteristics

A flow chart of study inclusion is presented in Figure 1. Database searches and review of prior meta-analyses yielded 2,117 studies, which were considered for inclusion in the meta-analysis. As depicted in the PRISMA diagram, 28 studies (n = 1,853) were identified for inclusion in the meta-analyses (Figure 1). Characteristics of included studies are described in Table 1.

Bias risk assessment/Study quality ratings

A visual summary of study quality ratings generated using RevMan (Review Manager, 2014), is displayed in Figure 2. Overall, 5/28 (18%) studies had low risk of bias on all five outcomes, 4/28 (14%) had low risk of bias on 4/5 outcomes, 2/28 (7%) had low risk of bias on 3/5 outcomes, 11/28 (39%) had low risk of bias on 2/5 outcomes, 5/28 (18%) had low risk of bias on 1/5 outcomes, and 2/28 (7%) did not have low risk of bias any of the five outcomes.

Among included studies, random sequence generation yielded a low risk of bias in 11/28 (39%) of studies and was unclear in 17/28 (61%) of studies. Allocation concealment yielded a low risk of bias in 7/28 (25%) of studies, was unclear in 19/28 (68%) of studies, and was high in 2/28 (7%) studies. Blinding of outcome assessment yielded a low risk of bias in 22/28 (79%) of studies, was unclear in 5/28 (18%) of studies, and was high in one (4%) study. Incomplete outcome data yielded a low risk of bias in 21/28 (75%) of studies, was unclear in 6/28 (21%) of studies, and was high in one (4%) study. Selective reporting yielded a low risk of bias in 8/28 (29%) of studies, was unclear in 19/28 (68%) of studies, and was high in one (4%) study.

Effect of behavioral activation and publication bias

A summary table of the primary outcome analyses is displayed in Table 2.

Effect of behavioral activation on depression compared to inactive control. BA for depression was compared to inactive controls using data from 19 studies and 844 participants.

The effect of BA versus inactive control was $g = 0.83$ (95% CI [0.58, 1.08], $p < 0.001$), representing a large effect size (Figure 4a). There was significant and moderate to high heterogeneity ($Q = 48.88$, $I^2 = 63$ (95% CI [40, 99]), $p < 0.001$), which is greater than what would be expected due to sampling error.

The funnel plot was visually inspected and revealed asymmetry, suggesting publication bias (Figure 3b). Tests of publication bias revealed a high likelihood of publication bias. The Duval and Tweedie trim and fill procedure imputed 7 studies (Duval and Tweedie adjusted $g = 0.47$ (95% CI [0.19, 0.75]). Results of the Egger's test also indicated a high likelihood of publication bias (Egger test intercept = 2.66, SE = 0.74, $p = 0.002$).

Moderation analysis. A subgroup analysis using a mixed-effects model was used to evaluate the moderating effect of the inclusion of a discussion on values in behavioral activation

compared to inactive control using data from 19 studies and 844 participants. The effect size for the subgroup with a discussion of values was $g = 0.86$ whereas the effect size for without values was $g = 0.82$. Using this model, presence of a discussion on values did not predict a significantly different depression effect size ($Q\ (1) = 0.07$, $p = 0.79$).

Effect of behavioral activation on depression compared to active control. BA for depression was compared to active controls using data from 15 studies and 1,098 participants. The effect of BA versus active control was $g = 0.15$ (95% CI -0.02, 0.33], $p = 0.084$), representing a negligible effect size (Figure 5a). There was significant and moderate heterogeneity ($Q = 23.73$, $I^2 = 41$ (95% CI [0, 98]), $p = 0.049$), which is greater than what would be expected due to sampling error.

The funnel plot was visually inspected and did not reveal substantial asymmetry, suggesting minimal publication bias (5b). Tests of publication bias revealed a low likelihood of publication bias. The Duval and Tweedie trim and fill procedure did not impute any studies or adjust the effect size. Results of the Egger's test also indicated a low likelihood of publication bias (Egger test intercept = 0.20, SE = 0.79, $p = 0.80$).

Effect of behavioral activation on anxiety compared to inactive control. BA for anxiety was compared to inactive controls using data from 5 studies and 426 participants. The effect of BA versus inactive control was $g = 0.37$ (95% CI [0.18, 0.57], $p < 0.001$), representing a small effect size (Figure 5a). There was not significant heterogeneity ($Q = 1.58$, $I^2 = 0$ (95% CI [0, 73]), $p = 0.81$), which is consistent with what would be expected due to sampling error. However this metric should be interpreted with caution due to the small number of studies included in this analysis.

The funnel plot was visually inspected and did not reveal substantial asymmetry, suggesting minimal publication bias (Figure 5b). Tests of publication bias revealed a low likelihood of publication bias. The Duval and Tweedie trim and fill procedure did not impute any studies or adjust the effect size. Results of the Egger's test also indicated a low likelihood of publication bias (Egger test intercept = -0.12, SE = 0.96, $p = 0.91$).

Effect of behavioral activation on anxiety compared to active control. BA for anxiety was compared to active controls using data from 4 studies and 599 participants. The effect of BA versus active control was $g = 0.03$ (95% CI [-0.13, 0.19], $p = 0.74$), representing a negligible effect (Figure 6a). There was no significant heterogeneity ($Q = 0.92$, $I^2 = 0$ (95% CI [0, 66], $p = 0.92$), which is consistent with what would be expected due to sampling error. However, this metric should be interpreted with caution due to the small number of studies included in this analysis.

The funnel plot was visually inspected and did not reveal substantial asymmetry, suggesting minimal publication bias (Figure 6b). Tests of publication bias revealed a low likelihood of publication bias The Duval and Tweedie trim and fill procedure did not impute any studies or adjust the effect size. Results of the Egger's test also indicated a low likelihood of publication bias (Egger test intercept = -0.11, SE = 0.90, $p = 0.92$).

Effect of behavioral activation on activation compared to inactive control. BA for activation was compared to inactive controls using data from 8 studies and 358 participants. The effect of BA versus inactive control was $g = 0.64$ (95% CI [0.39, 0.88], $p < 0.001$), representing a medium effect size (Figure 7a). There was no significant heterogeneity ($Q = 8.58$, $I^2 = 18$ (95% CI [0, 96], p = 0.29), which is consistent with what would be expected due to sampling error.

However, this metric should be interpreted with caution due to the small number of studies included in this analysis.

The funnel plot was visually inspected and revealed slight asymmetry, suggesting potential publication bias (Figure 7b). Tests of publication bias revealed a moderate likelihood of publication bias. The Duval and Tweedie trim and fill procedure imputed 1 study (Duval and Tweedie adjusted $g = 0.60$ (95% CI [0.34, 0.85]). However, the Egger's test did not find a high likelihood of publication bias (Egger test intercept = 0.73, SE = 1.22, $p = 0.57$).

Effect of behavioral activation on activation compared to active control. BA for activation was compared to active controls using data from 4 studies and 157 participants. The effect of BA versus active control was $g = 0.04$ (95% CI [-0.27, 0.35], $p = 0.80$) (Figure 8a), representing a negligible effect. There was no significant heterogeneity ($Q = 0.31$, $I^2 < 0.01$ (95% CI [0, 52], $p = 0.96$), which is consistent with what would be expected due to sampling error. However, this metric should be interpreted with caution due to the small number of studies included in this analysis.

The funnel plot was visually inspected and did not reveal substantial asymmetry, suggesting minimal publication bias (Figure 8b). Tests of publication bias revealed a low likelihood of publication bias The Duval and Tweedie trim and fill procedure did not impute any studies or adjust the effect size. Results of the Egger's test also indicated a low likelihood of publication bias (Egger test intercept = 0.31, SE = 0.63, $p = 0.67$).

Discussion

In addition to depressive symptoms, changes in anxiety symptoms and activation are important outcomes in behavioral activation. Thus, in the present study, we aimed to update and build on the extant literature documenting the antidepressant effects of behavioral activation by

analyzing the efficacy BA for depression, anxiety, and activation. We also examined the moderating effect of a discussion of values on depression symptoms.

Results from studies comparing BA to inactive control yielded a large effect size (g = 0.83) for reducing depression symptoms. Results from studies comparing BA to active control conditions yielded a small, non-significant effect size (g = 0.15). These findings are consistent with prior research examining the effect of BA on depressive symptoms relative to inactive control conditions (g = 0.74 in Ekers 2014). However, the present study examined a larger body of literature. Despite this, we observed a large effect relative to inactive control, and only one study with findings favoring the inactive control condition over BA.

Results from studies comparing BA to inactive control yielded a small effect size (g=0.37, p<0.001) for reducing anxiety symptoms. Results from studies comparing BA to active control conditions yielded a negligible, nonsignificant effect size (g=0.03). To our knowledge, these results provide the first pooled estimates of controlled effect sizes of behavioral activation for improving anxiety symptoms. This effect size is somewhat smaller than what has been observed in a prior meta-analysis of psychotherapy for depression trials (g = 0.52; (Weitz *et al.* 2018). Transdiagnostic treatment protocols, such as the Unified Protocol (Farchione *et al.* 2012) or Acceptance and Commitment Therapy (Hayes *et al.* 1999) may offer a more parsimonious option for addressing comorbidity. Interestingly, however, there has been mixed evidence regarding combining depression and anxiety focused CBT protocols, which do not appear to outperform CBT for depression (Shafran *et al.* 2018).

Results from studies comparing BA to inactive control yielded a medium effect (g=0.64, p<0.001) size for increasing activation. Results from studies comparing BA to active control conditions yielded a small effect size (g=0.04). The present study is the first meta-analysis to

report a pooled effect for improving activation in BA. Importantly, these findings suggest that behavioral activation is effectively engaging a putative mechanistic target of BA. It is interesting that this effect was not larger in BA than in other interventions. This may point to the bi-directional nature of activation and depressed mood. According to the behavioral model of depression, one would expect an increase in activation to precede improvement in depressed mood, but would also expect that an improvement in mood would result in increased activation. So, it is possible that any treatment that is effective in improving depressed mood will also increase activation. Thus, these findings provide preliminary support for the bidirectional pathway of activation and mood in the behavioral model of depression. This is consistent with prior research investigating other mechanisms of change within treatment for depression. For example, a meta-analysis of dysfunctional thinking in CBT did not find a significant difference between CBT and other psychotherapies or pharmacotherapies for reducing dysfunctional thinking (Cristea *et al.* 2015). Collectively, these findings highlight the need for further research understanding the working mechanism of psychological interventions for depression (Cuijpers *et al.* 2019c).

Nevertheless, the moderate magnitude of the effect size suggests that there may be room for improvement in increasing activation in BA. One viable target for increasing activation may be through improving homework compliance. For example, a recent study demonstrated higher compliance rates following a brief guided practice during a single session intervention (Stein *et al.* 2019). Another approach to increasing activation is Positive Affect Training (PAT), which incorporates a number of strategies for increasing reward sensitivity to increase positive affect (Craske *et al.* 2019). Of note, this treatment included a module termed “Augmented Behavioral Activation Training” which employs novel strategies for enhancing reinforcement. In addition to

activity planning, the augmented BA includes therapist guided recounting exercises in which the patient is asked to imagine and focus on positive affective experiences. This is hypothesized to provide additional reinforcement to the patient and may be another viable augmentation strategy for BA.

We also explored the association between the presence of a discussion of values in treatment with the effect size of the intervention. We did not find a significant effect of values on depression treatment outcome. Although it is possible that there is no added benefit of including a values discussion, it is also possible that we did not find an effect because of the small number of studies that examined this variable. There were only 6 studies that included a values discussion included in our analyses and three of them were only single session interventions. We were underpowered to include additional covariates in this model, but future research with a larger number of studies, should control for the number of sessions. Nonetheless, additional research in the field is needed directly comparing BA with and without a values discussion in a well-powered randomized controlled trial to better address this question.

These results must be considered in light of several limitations. First, It is important to note that the confidence intervals around these effect sizes were large and many studies were of low quality and/or had small sample sizes, especially the analyses of anxiety, activation, and values. Although it is possible to examine this empirically by conducting sensitivity analyses using studies with low risk of bias, we did not have a sufficient number of low risk studies to perform meaningful sensitivity analyses (i.e. fewer than five low risk studies per comparison).Second, there was considerable variability in the "dosing" of the intervention, with the number of sessions ranging from 1 to 24, however previous meta-analyses have found that the number of sessions was not a significant moderator of treatment effect (Ekers *et al.* 2014).

Third, several studies included multiple active comparison condition, which resulted in the exclusion of some data from the analysis. Fourth, there was considerable evidence of publication bias in the included studies, which suggests that there may have been negative trials of BA that were not published. This implies that the true effect size of behavioral activation relative to control may be smaller than the observed effect size. Finally, there were too few studies measuring the long-term effects of BA, so we were only able to examine the acute effects of the intervention. In addition to more studies with larger samples, many of the included studies are older and were published before reporting guidelines, such as PRISMA existed. Thus, the risk of bias in many studies is high. There is a need for well-powered, low risk of bias (i.e. high quality) studies of BA.

Clinically, these results suggest that BA alone may be an appropriate treatment selection when individuals present with a primary concern of depression and co-occurring symptoms of anxiety. However, for individuals who place equal concern on depression and anxiety symptoms or who are experiencing high levels of anxiety, clinicians may consider a range of alternative treatment approaches. Given the brevity and flexibility of behavioral activation protocols, clinicians may consider completing a brief course of BA with patients prior to initiating targeted treatment for anxiety.

In summary, our meta-analysis provides updated evidence supporting the efficacy of behavioral activation for depression, anxiety, and activation. These findings support BA as at least as effective as other active treatments for depression and substantially more effective than inactive control conditions for all outcomes. We did not find an effect of including a discussion on values on depression symptom outcome. Additional well-powered studies of behavioral activation are needed, particularly comparing different versions of the intervention. Clinically,

these result support the use of behavioral activation for depression across a variety of populations and settings, including for individuals with co-occurring anxiety symptoms. Future versions of BA may consider augmentation strategies for bolstering the effects of the intervention on activation and anxiety.

References

Armento MEA, McNulty JK, Hopko DR (2012). Behavioral activation of religious behaviors (BARB): Randomized trial with depressed college students. *Psychology of Religion and Spirituality* **4**, 206–222.

Barth J, Munder T, Gerger H, Nüesch E, Trelle S, Znoj H, Jüni P, Cuijpers P (2013). Comparative Efficacy of Seven Psychotherapeutic Interventions for Patients with Depression: A Network Meta-Analysis. ed. A. C. Tsai *PLoS Medicine* **10**, e1001454.

Bolton P, Bass JK, Zangana GAS, Kamal T, Murray SM, Kaysen D, Lejuez CW, Lindgren K, Pagoto S, Murray LK, Van Wyk SS, Ahmed AMA, Mohammad Amin NM, Rosenblum M (2014). A randomized controlled trial of mental health interventions for survivors of systematic violence in Kurdistan, Northern Iraq. *BMC Psychiatry* **14**, 360.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2011). Introduction to meta-analysis. John Wiley & Sons.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. Comprehensive meta-analysis (version 3). Biostat, Englewood, NJ (2013)

Collado A, Calderón M, MacPherson L, Lejuez C (2016). The efficacy of Behavioral Activation Treatment among Depressed Spanish-speaking Latinos. *Journal of consulting and clinical psychology* **84**, 651–657.

Covidence systematic review software, Veritas Health Innovation, Melbourne, Australia. (2019) Retrieved from www.covidence.org.

Craske MG, Meuret AE, Ritz T, Treanor M, Dour H, Rosenfield D (2019). Positive affect treatment for depression and anxiety: A randomized clinical trial for a core feature of anhedonia. *Journal of Consulting and Clinical Psychology* **87**, 457–471.

Cristea IA, Huibers MJH, David D, Hollon SD, Andersson G, Cuijpers P (2015). The effects of cognitive behavior therapy for adult depression on dysfunctional thinking: A meta-analysis. *Clinical Psychology Review* **42**, 62–71.

Cuijpers, P., Karyotaki, E., & Ciharova, M. (2019a, August 30). A meta-analytic database of randomised trials on psychotherapies for depression. https://doi.org/10.17605/OSF.IO/825C6

Cuijpers P, Karyotaki E, de Wit L, Ebert DD (2019b). The effects of fifteen evidence-supported therapies for adult depression: A meta-analytic review. *Psychotherapy Research*, 1–15.

Cuijpers P, Reijnders M, Huibers MJH (2019c). The Role of Common Factors in Psychotherapy Outcomes. *Annual Review of Clinical Psychology* **15**, 207–231.

Cuijpers P, van Straten A, Warmerdam L (2007). Behavioral activation treatments of depression: a meta-analysis. *Clinical Psychology Review* **27**, 318–326.

Cullen JM, Spates CR, Pagoto S, Doran N (2006). Behavioral activation treatment for major depressive disorder: A pilot investigation. *The Behavior Analyst Today* **7**, 151–166.

Delgadillo J, Gore S, Ali S, Ekers D, Gilbody S, Gilchrist G, McMillan D, Hughes E (2015). Feasibility Randomized Controlled Trial of Cognitive and Behavioral Interventions for Depression Symptoms in Patients Accessing Drug and Alcohol Treatment. *Journal of Substance Abuse Treatment* **55**, 6–14.

Dimidjian S, Goodman SH, Sherwood NE, Simon GE, Ludman E, Gallop R, Welch SS, Boggs JM, Metcalf CA, Hubley S, Powers JD, Beck A (2017). A pragmatic randomized clinical trial of behavioral activation for depressed pregnant women. *Journal of Consulting and Clinical Psychology* **85**, 26–36.

Dimidjian S, Hollon SD, Dobson KS, Schmaling KB, Kohlenberg RJ, Addis ME, Gallop R, McGlinchey JB, Markley DK, Gollan JK, Atkins DC, Dunner DL, Jacobson NS (2006). Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. *Journal of Consulting and Clinical Psychology* **74**, 658–670.

Duval S, Tweedie R (2000). Trim and Fill: A Simple Funnel-Plot-Based Method of Testing and Adjusting for Publication Bias in Meta-Analysis. *Biometrics* **56**, 455–463.

Egger M, Smith GD, Schneider M, Minder C (1997). Bias in meta-analysis detected by a simple, graphical test. *Bmj* **315**, 629–634.

Ekers D, Richards D, McMillan D, Bland JM, Gilbody S (2011). Behavioural activation delivered by the non-specialist: phase II randomised controlled trial. *British Journal of Psychiatry* **198**, 66–72.

Ekers D, Webster L, Van Straten A, Cuijpers P, Richards D, Gilbody S (2014). Behavioural Activation for Depression; An Update of Meta-Analysis of Effectiveness and Sub Group Analysis. ed. A. Aleman *PLoS ONE* **9**, e100100.

Farchione TJ, Fairholme CP, Ellard KK, Boisseau CL, Thompson-Hollands J, Carl JR, Gallagher MW, Barlow DH (2012). Unified Protocol for Transdiagnostic Treatment of Emotional Disorders: A Randomized Controlled Trial. *Behavior Therapy* **43**, 666–678.

Ferster CB (1973). A functional analysis of depression. *American psychologist* **28**, 857.

Gawrysiak M, Nicholas C, Hopko DR (2009). Behavioral activation for moderately depressed university students: Randomized controlled trial. *Journal of Counseling Psychology* **56**, 468–475.

Hayes SC, Strosahl KD, Wilson KG (1999). *Acceptance and commitment therapy: An experiential approach to behavior change*. Acceptance and commitment therapy: An experiential approach to behavior change. Guilford Press: New York, NY, US.

Hedges LV, Olkin I (2014). *Statistical Methods for Meta-Analysis*. Academic Press.

Higgins JP, Thompson SG, Deeks JJ, Altman DG (2003). Measuring inconsistency in meta-analyses. *Bmj* **327**, 557–560.

Higgins JPT, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, Savovic J, Schulz KF, Weeks L, Sterne JAC, Cochrane Bias Methods Group, Cochrane Statistical Methods Group (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* **343**, d5928–d5928.

Hopko DR, Armento MEA, Robertson SMC, Ryba MM, Carvalho JP, Colman LK, Mullane CN, Gawrysiak MJ, Bell JL, Mcnulty JK, Lejuez CW (2011). Brief behavioral activation and problem-solving therapy for depressed breast cancer patients: randomized trial. *Journal of consulting and clinical psychology* **79**, 834–849.

Hopko DR, Lejuez CW, Lepage JP, Hopko SD, McNeil DW (2003). A Brief Behavioral Activation Treatment for Depression: A Randomized Pilot Trial within an Inpatient Psychiatric Hospital. *Behavior Modification* **27**, 458–469.

Jacobson NS, Dobson KS, Truax PA, Addis ME, Koerner K, Gollan JK, Gortner E, Prince SE (1996). A component analysis of cognitive-behavioral treatment for depression. *Journal of Consulting and Clinical Psychology* **64**, 295–304.

Jahoda A, Hastings R, Hatton C, Cooper S-A, Dagnan D, Zhang R, McConnachie A, McMeekin N, Appleton K, Jones R, Scott K, Fulton L, Knight R, Knowles D, Williams C, Briggs A, MacMahon K, Lynn H, Smith I, Thomas G, Melville C (2017). Comparison of behavioural activation with guided self-help for treatment of depression in adults with intellectual disabilities: a randomised controlled trial. *The Lancet Psychiatry* **4**, 909–919.

Kanter JW, Santiago-Rivera AL, Santos MM, Nagy G, López M, Hurtado GD, West P (2015). A Randomized Hybrid Efficacy and Effectiveness Trial of Behavioral Activation for Latinos With Depression. *Behavior Therapy* **46**, 177–192.

Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry* **62**, 593–602.

Lamers, F., van Oppen, P., Comijs, H.C., Smit, J.H., Spinhoven P, van Balkom AJ, Nolen WA, Zitman FG, Beekman, A.T.F., Penninx BWJH (2011). Comorbidity patterns of anxiety and depressive disorders in a large cohort study: the Netherlands Study of Depression and Anxiety (NESDA). *Journal of Clinical Psychiatry* **72**, 341–348.

Lejuez CW, Hopko DR, Acierno R, Daughters SB, Pagoto SL (2011). Ten Year Revision of the Brief Behavioral Activation Treatment for Depression: Revised Treatment Manual. *Behavior Modification* **35**, 111–161.

Lejuez CW, Hopko DR, Hopko SD (2001). A Brief Behavioral Activation Treatment for Depression: Treatment Manual. *Behavior Modification* **25**, 255–286.

Lewinsohn PM, Sullivan JM, Grosscup SJ (1980). Changing reinforcing events: an approach to the treatment of depression. *Psychotherapy: Theory, Research & Practice* **17**, 322.

MacPhillamy DJ, Lewinsohn PM (1974). Depression as a function of levels of desired and obtained pleasure. *Journal of Abnormal Psychology* **83**, 651–657.

Martell CR, Dimidjian S, Herman-Dunn R (2013). *Behavioral activation for depression: A clinician's guide*. Guilford Press.

Mazzucchelli T, Kane R, Rees C (2009). Behavioral Activation Treatments for Depression in Adults: A Meta-analysis and Review. *Clinical Psychology: Science and Practice* **16**, 383–411.

McIndoo CC, File AA, Preddy T, Clark CG, Hopko DR (2016). Mindfulness-based therapy and behavioral activation: A randomized controlled trial with depressed college students. *Behaviour Research and Therapy* **77**, 118–128.

McNamara K, Horan JJ (1986). Experimental construct validity in the evaluation of cognitive and behavioral treatments for depression. *Journal of Counseling Psychology* **33**, 23–30.

Meeks S, Looney SW, Van Haitsma K, Teri L (2008). BE-ACTIV: A Staff-Assisted Behavioral Intervention for Depression in Nursing Homes. *The Gerontologist* **48**, 105–114.

Moher D, Liberati A, Tetzlaff J, Altman DG, Group TP (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLOS Medicine* **6**, e1000097.

Moradveisi L, Huibers MJH, Renner F, Arasteh M, Arntz A (2013). Behavioural activation v. antidepressant medication for treating depression in Iran: randomised trial. *British Journal of Psychiatry* **202**, 204–211.

Myhre MØ, Strømgren B, Arnesen EF, Veland MC (2018). *The feasibility of brief behavioural activation treatment for depression in a PICU: a systematic replication*. Text

Nasrin F, Rimes K, Reinecke A, Rinck M, Barnhofer T (2017). Effects of brief behavioural activation on approach and avoidance tendencies in acute depression: Preliminary findings. *Behavioural and cognitive psychotherapy* **45**, 58–72.

Padfield M (1976). The comparative effects of two counseling approaches on the intensity of depression among rural women of low socioeconomic status. *Journal of Counseling Psychology* **23**, 209–214.

Review Manager (RevMan) [Computer program]. Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Richards DA, Ekers D, McMillan D, Taylor RS, Byford S, Warren FC, Barrett B, Farrand PA, Gilbody S, Kuyken W, O'Mahen H, Watkins ER, Wright KA, Hollon SD, Reed N, Rhodes S, Fletcher E, Finning K (2016). Cost and Outcome of Behavioural Activation versus Cognitive Behavioural Therapy for Depression (COBRA): a randomised, controlled, non-inferiority trial. *Lancet (London, England)* **388**, 871–880.

Shafran R, Wroe A, Nagra S, Pissaridou E, Coughtrey A (2018). Cognitive behaviour treatment of co-occurring depression and generalised anxiety in routine clinical practice. ed. I. Branchi *PLOS ONE* **13**, e0201226.

Snarski M, Scogin F, DiNapoli E, Presnell A, McAlpine J, Marcinak J (2011). The Effects of Behavioral Activation Therapy With Inpatient Geriatric Psychiatry Patients. *Behavior Therapy* **42**, 100–108.

Stein, A., Shumake, J., Beevers, C. G., & Smits, J. A. J. (2019, September 27). Therapist Guided Activity Practice for Depressive Symptoms in University Students: A Randomized Controlled Trial. https://doi.org/10.31234/osf.io/4u7qx

Stein, A.T., Smits, J.A.J., Carl, E. Emily Carl. (n.d.). Behavioral activation for depression: a meta-analysis. PROSPERO 2019 CRD42019124300. Retrieved from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42019124300

Sturmey P (2009). Behavioral Activation Is an Evidence-Based Treatment for Depression. *Behavior Modification* **33**, 818–829.

Taylor FG, Marshall WL (1977). Experimental analysis of a cognitive-behavioral therapy for depression. *Cognitive Therapy and Research* **1**, 59–72.

Thompson LW, Gallagher D, Breckenridge JS (1987). Comparative effectiveness of psychotherapies for depressed elders. *Journal of Consulting and Clinical Psychology* **55**, 385.

Travers C (2017). Increasing enjoyable activities to treat depression in nursing home residents with dementia: A pilot study. *Dementia* **16**, 204–218.

Weitz E, Kleiboer A, van Straten A, Cuijpers P (2018). The effects of psychotherapy for depression on anxiety symptoms: a meta-analysis. *Psychological Medicine* **48**, 2140–2152.

Wilson PH, Goldin JC, Charbonneau-Powis M (1983). Comparative efficacy of behavioral and cognitive treatments of depression. *Cognitive Therapy and Research* **7**, 111–124.

World Health Organization (2017). *Depression and other common mental disorders: global health estimates*. World Health Organization.

Yokoyama S, Okamoto Y, Takagaki K, Okada G, Takamura M, Mori A, Shiota S, Ichikawa N, Jinnin R, Yamawaki S (2018). Effects of behavioral activation on default mode network connectivity in subthreshold depression: A preliminary resting-state fMRI study. *Journal of Affective Disorders* **227**, 156–163.

Table 1. Properties of included studies

Study First Author	Year	Conditions	Control Type	N by group	Values	No. of sessions*	Sample population	Country	Outcome measure(s)
Armento	2012	SC	Inactive	SC = 25 BA = 25	BA	1	University	USA	D = BDI-II Ax = BAI A = EROS
Bolton	2014	WL	Inactive	WL = 66 BA = 114	BA	12	Survivors of systematic violence	Iraq	D = HSCL-D Ax = HSCL-A
Collado	2016	SC	Inactive	SC = 12 BA = 15	BA	10	Community	USA	D = BDI-II A = BADS/RPI
Cullen	2006	WL	Inactive	WL = 12 BA = 13	Neither	6	Community	USA	D = HRSD/BDI-II
Delgadillo	2015	GSH	Active	GSH = 20 BA = 19	Neither	12	Alcohol or drug dependence	UK	D = PHQ-9
Dimidjian	2006	PL CT	Inactive Active	PL = 41 CT = 35 BA = 29	Neither	24	Community	USA	D = BDI-II/HRSD
Dimidjian	2017	TAU	Inactive	TAU =68 BA= 70	Neither	10	Pregnant women	USA	D = PHQ-9 Ax = GAD-7 A = EROS
Ekers	2011	TAU	Inactive	TAU = 22 BA = 16	Neither	12	Primary care	UK	D = BDI-II
Gawrysiak	2009	WL	Inactive	WL = 16 BA = 14	BA	1	University	USA	D = BDI-II Ax = BAI A = EROS
Hopko	2003	SC	Inactive	SC = 15 BA = 10	BA	6	Inpatient	USA	D = BDI-II
Hopko	2011	PST	Active	PST = 38 BA = 42	BA	8	Breast cancer	USA	D = BDI-II/HRSD Ax = BAI A = EROS
Jacobson	1996	CT	Active	CT = 43 BA = 56	Neither	20	Community	USA	D = BDI-II/HRSD
Jahoda	2018	GSH	Active	GSH = 75 BA = 74	BA	12	Intellectual Disabilities	UK	D = GDS-LD Ax = GAS-ID

Kanter	2015	TAU	Active	TAU = 22 BA = 21	BA	12	Latinx	USA	D = BDI-II/HRSD
McIndoo	2016	Mindfulness	Active	M = 18 WL = 13 BA= 15	BA	4	University	USA	D = BDI-II/HRSD Ax = BAI
McNamara	1986	CT	Active	CT = 10 SC = 10 BA = 10	Neither	8	University	USA	D = BDI A = PES
Meeks	2008	TAU	Inactive	TAU = 4 BA = 10	Neither	10	Nursing home	USA	D = GDS/HRSD
Moradveisi	2013	ADM	Active	ADM = 35 BA = 45	Neither	16	Community	Iran	D = BDI-II/HRSD
Myhre	2018	TAU	Active	TAU = 10 BA = 9	Both	M=4.84	Clinical	Norway	D = BDI-II
Nasrin	2017	WL	Inactive	WL = 17 BA = 19	BA	1	Clinical	UK	D = PHQ-9 A = BADS
Padfield	1976	SC	Inactive	SC = 12 BA = 12	Neither	12	Low SES	USA	D = Zung
Richards	2016	CBT	Active	CBT = 189 BA = 175	Neither	M=11.5	Primary Care	UK	D = PHQ-9 Ax = GAD-7
Snarski	2011	TAU	Active	TAU = 19 BA = 21	Neither	8	Inpatient geriatric	USA	D = GDS-S
Taylor	1977	CT WL	Active Inactive	CT = 7 WL = 7 BA = 7	Neither	6	University	Canada	D = BDI
Thompson	1987	PDN WL	Active Inactive	PDN = 20 WL = 19 BA= 21	Neither	18	Community older adults	USA	D = BDI/HRSD/GDS A = PES
Travers	2017	WAT	Inactive	WAT = 6 BA = 10	Neither	8	Nursing home	Australia	D = GDS-12
Wilson	1983	CT WL	Active Inactive	CT = 8 WL = 9 BA = 8	Neither	8	Community	Australia	D = BDI/HRSD A = PES
Yokoyama	2018	No TX	Inactive	NoTX = 21 BA = 19	Neither	5	University	Japan	D = BDI-II A= BADS-AC/EROS

Note. A= Activation, Ax=Anxiety, ADM = Antidepressant medication, BA = Behavioral activation, BADS = Behavioral Activation for Depression Scale, BAI = Beck Anxiety Inventory, BDI= Beck Depression Inventory, CT= Cognitive therapy, CBT = Cognitive behavioral therapy, D=Depression, EROS= Environmental Reward Observation Scale, GAD=Generalized Anxiety Disorder, GAD, GAS-ID= Glasgow Anxiety Scale for people with an Intellectual Disability, GDS = Geriatric Depression Scale, GDS-LD = Glasgow Depression Scale for people with a learning disability, GSH = Guided self-help, HRSD= Hamilton Rating Scale for Depression, HSCL= Hopkins Symptom Checklist, M = Mean number of completed sessions, PES=Pleasant events schedule, PHQ-9 = Patient Health Questionnaire-9, PDN = Psychodynamic, PL = Placebo, PST= Problem Solving Therapy, RPI=Reward Probability Index, SC= supportive counseling, TAU= treatment as usual, USA= United States of American, UK = United Kingdom, Zung = Zung self-rating depression scale, WAT=Walking and talking intervention, WL = waitlist
* Numbers of sessions represent the number of planned sessions, except for two studies which reported the mean number of completed sessions, denoted by "M=".

Table 2. Summary of primary outcomes

				Effect size			*Heterogeneity*			
Outcome	Control	$N_{Studies}$	$N_{Participants}$	*g*	[95% CI]	*p*	*q*	I^2	[95% CI]	*p*
Depression	Inactive	19	844	0.83	[0.58, 1.08]	< 0.01	48.88	63	[40, 99]	<0.01
Depression	Active	15	1,098	0.15	[-0.02, 0.33]	0.08	23.73	41	[0, 98]	0.05
Anxiety	Inactive	5	426	0.37	[0.18, 0.57]	< 0.01	1.58	0	[0, 73]	0.81
Anxiety	Active	4	599	0.03	[-0.13, 0.19]	0.74	0.92	0	[0, 66]	0.92
Activation	Inactive	8	358	0.64	[0.39, 0.88]	<0.01	8.58	18	[0, 96]	0.29
Activation	Active	4	157	0.04	[-0.27, 0.35]	0.08	0.31	0	[0, 52]	0.96

Figure 1. Flow diagram of study inclusion

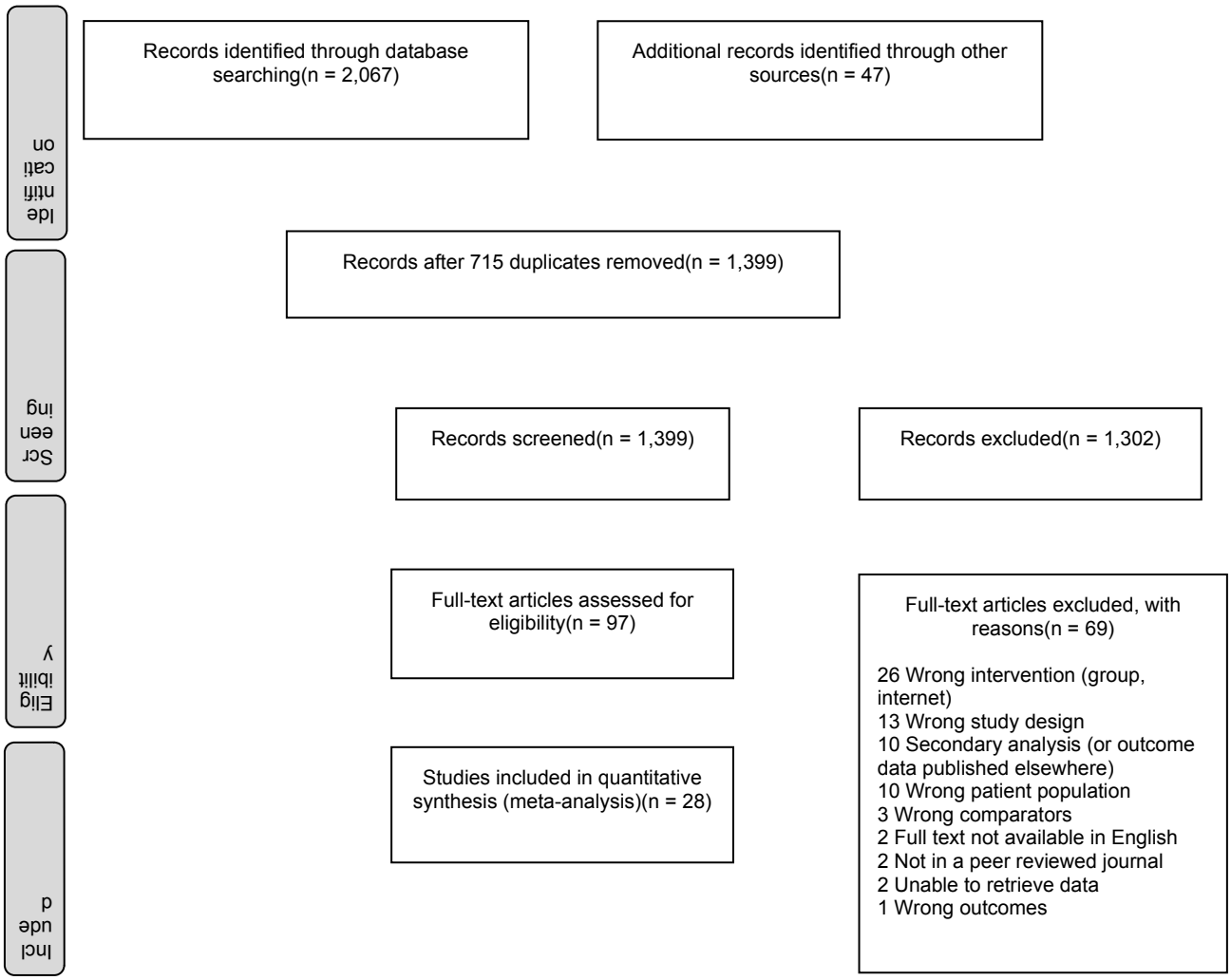

Figure 2. Study quality assessment

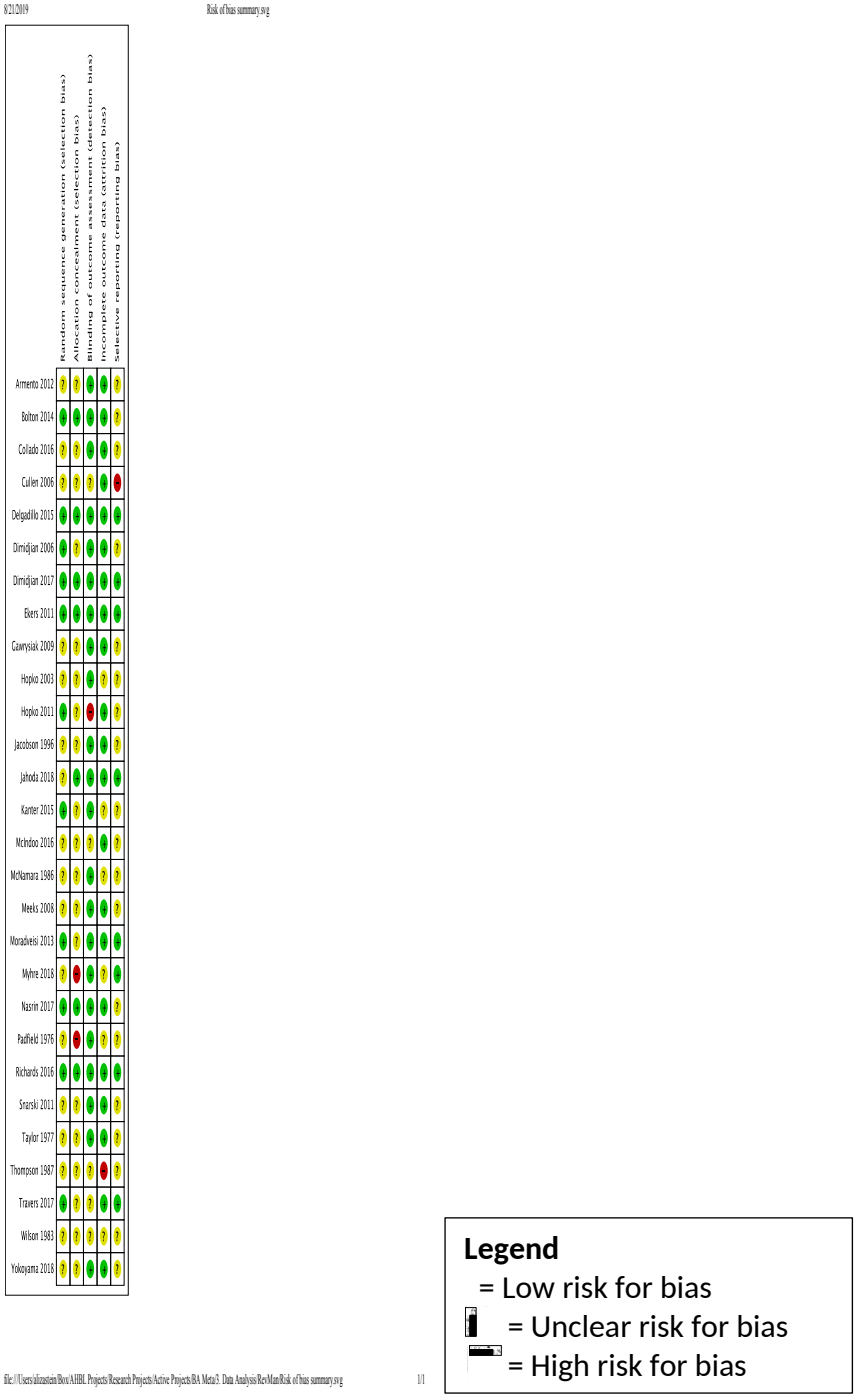

Figure 3a. Forrest Plot of Behavioral Activation vs Inactive Control for Depression Symptoms

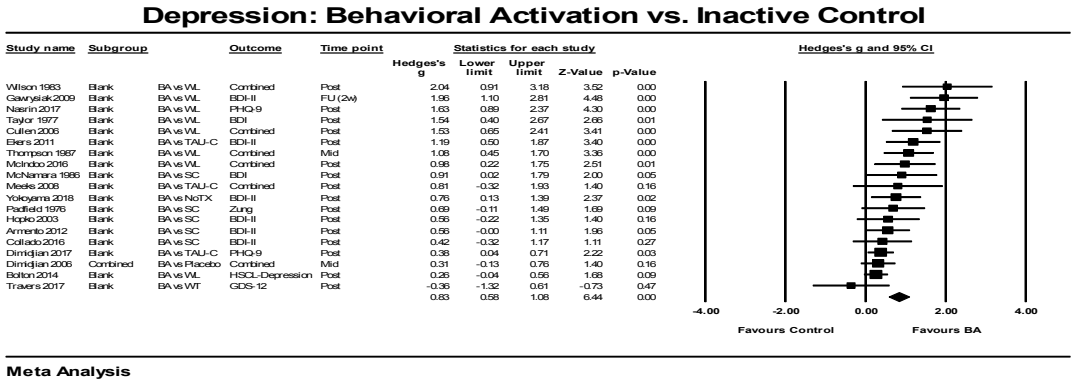

Figure 3a. Funnel Plot of Behavioral Activation vs Inactive Control for Depression Symptoms

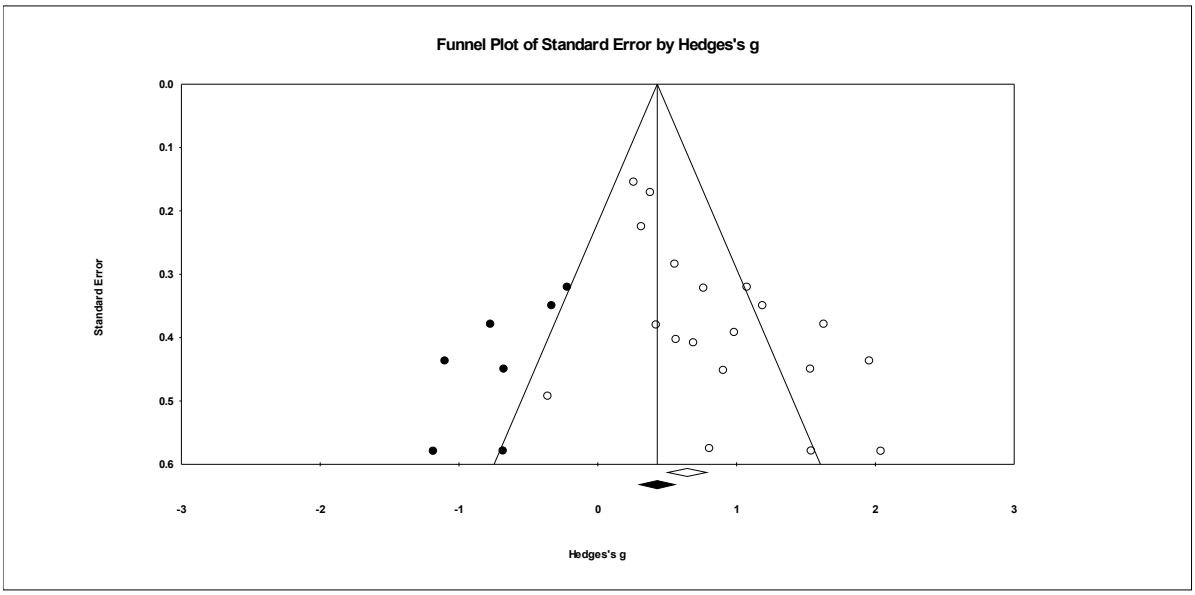

Figure 4a. Forrest Plot of Behavioral Activation vs Active Control for Depression Symptoms

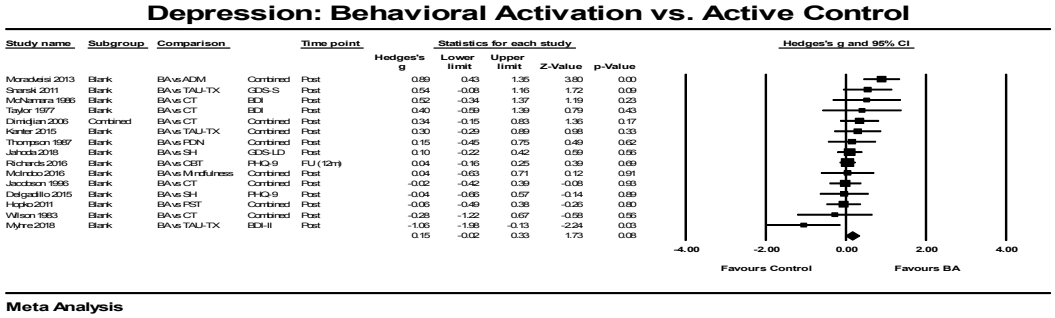

Figure 4b. Forrest Plot of Behavioral Activation vs Active Control for Depression Symptoms

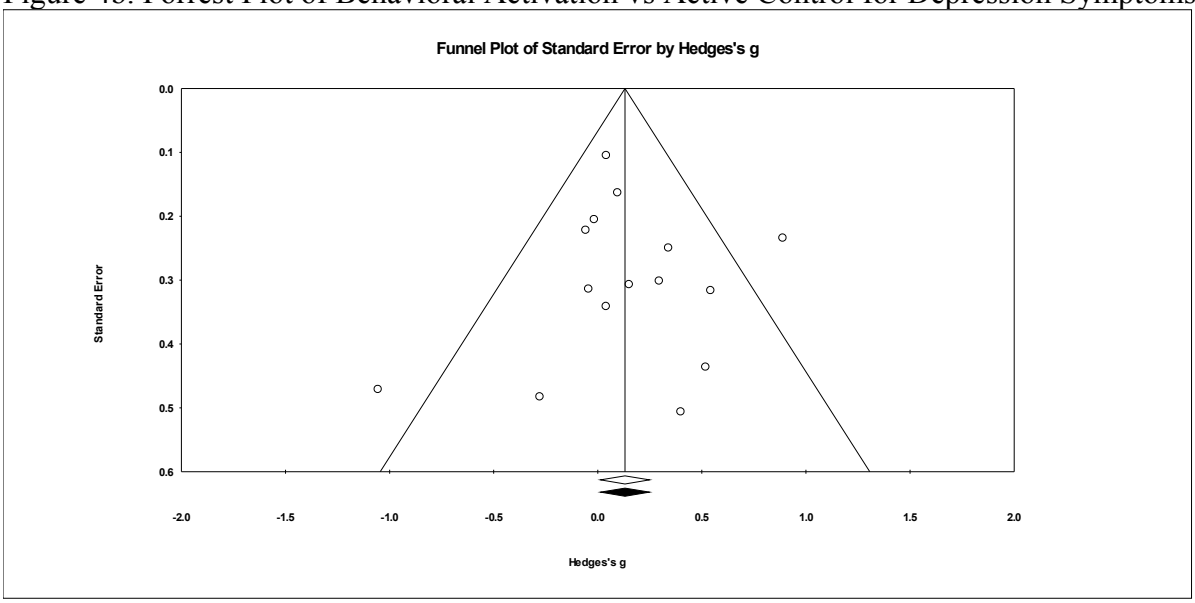

Figure 5a. Behavioral Activation vs Inactive Control for Anxiety Symptoms

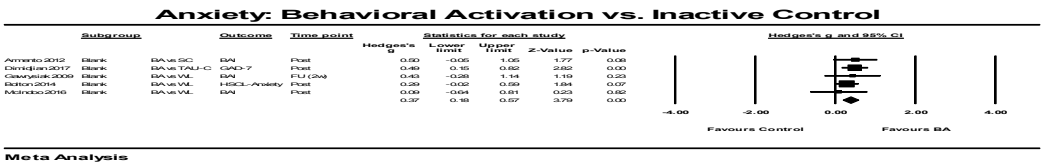

Figure 5b. Funnel plot for Behavioral Activation vs Inactive Control for Anxiety Symptoms

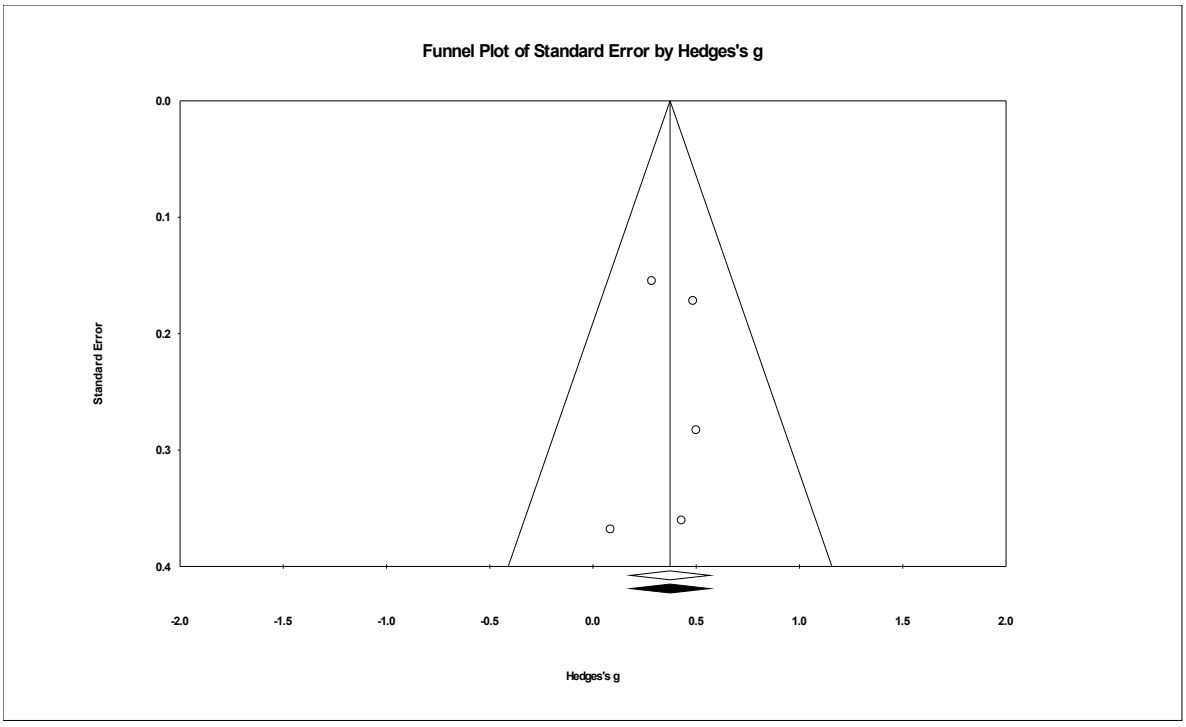

Figure 6a. Forrest Plot of Behavioral Activation vs Active Control for Anxiety Symptoms

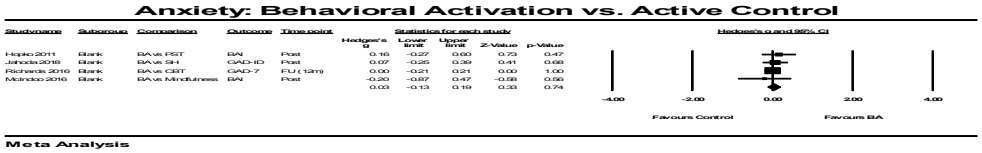

Figure 6b. Forrest Plot of Behavioral Activation vs Active Control for Anxiety Symptoms

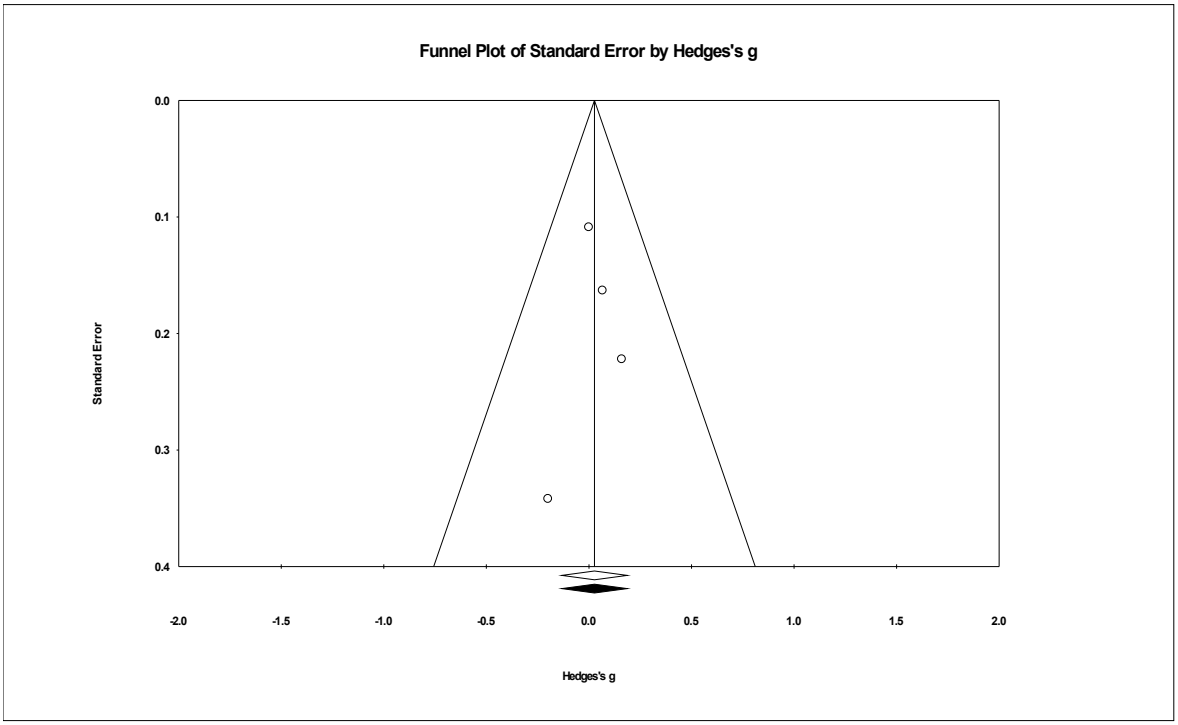

Figure 7a. Forrest Plot of Behavioral Activation vs Active Control for Activation

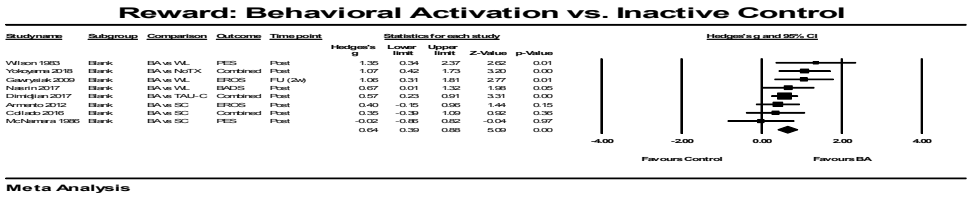

Figure 7b. Funnel Plot of Behavioral Activation vs Active Control for Activation

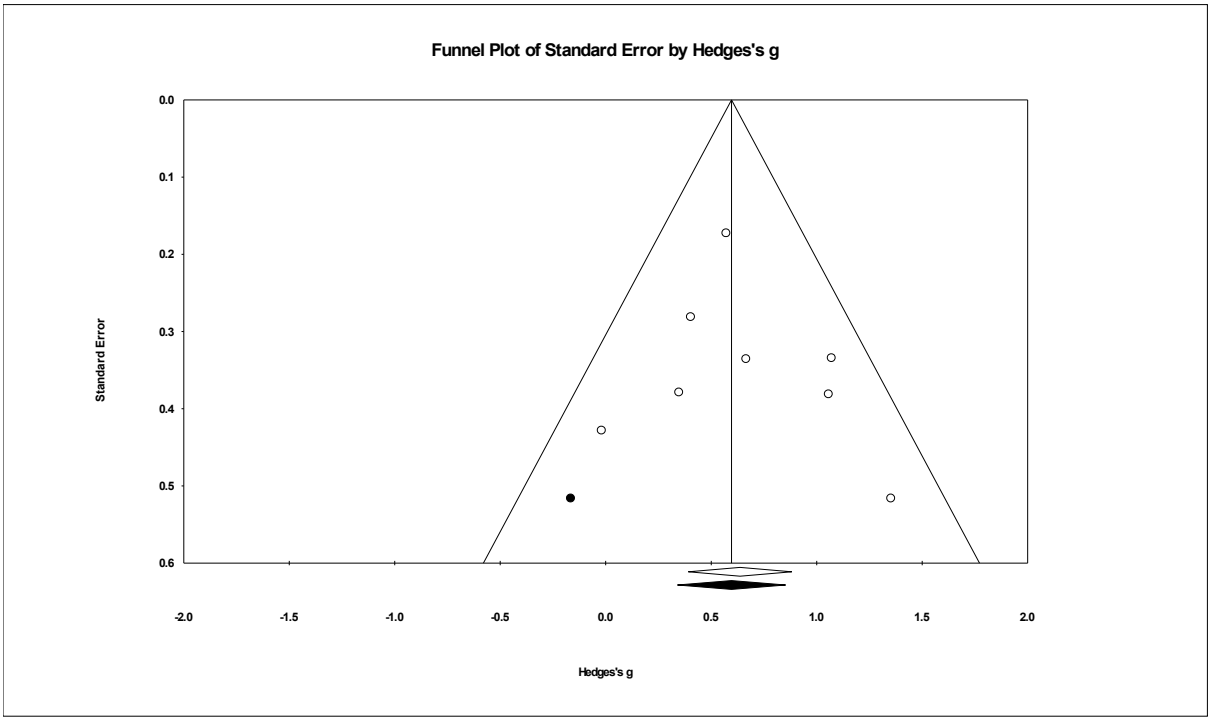

Figure 8a. Forrest Plot of Behavioral Activation vs Active Control for Activation

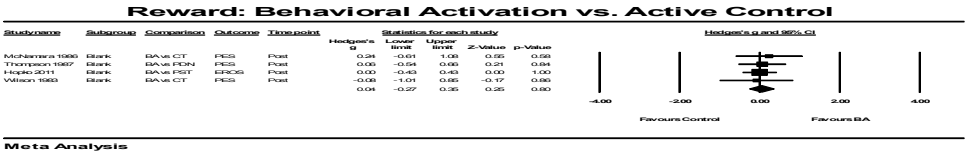

Figure 8b. Forrest Plot of Behavioral Activation vs Active Control for Activation