1   A Meta-Analysis of Expressive Writing on Quality of Life, Posttraumatic Growth, and

2   Posttraumatic Stress

Abstract

Emotional expression has been shown to be beneficial for promoting both positive psychological and physical health outcomes. Unfortunately, inhibiting emotions can lead to impairments in physical and psychological health. James Pennebaker showed that expressive writing is an effective form of emotional expression, and he and others have used expressive writing as an experimental manipulation to gauge its efficacy in treating a wide variety of health-related and psychological outcomes. While many studies have been conducted that examine the efficacy of expressive writing across such outcomes, a considerable amount of these studies tend to neglect necessary considerations such as power and meaningfulness of respective effect sizes. Six previous meta-analyses have been conducted that examine expressive writing's effect on psychological outcomes. However, these studies focus on the experimental versus control group effect size. Thus, our meta-analysis sought to examine the efficacy of an expressive writing task on only the experimental conditions in studies measuring posttraumatic growth, posttraumatic stress, and quality of life using random effects models. Results indicated a small overall effect size for posttraumatic stress and negligible to small effect sizes for posttraumatic growth and quality of life. Implications for future research design and interpretation of published research are discussed.

Keywords: meta-analysis, positive psychology, expressive writing

A Meta-Analysis of Expressive Writing on Quality of Life, Posttraumatic Growth, and

Posttraumatic Stress

Emotional Expression

Emotional expression relating to negative emotions or trauma has been shown to enhance both mental and physical health outcomes (Esterling, Antoni, Kumar, & Schneiderman, 1990; Fawzy et al., 1993; Lieberman & Goldstein, 2006; Rachman, 1980; Scheff, 1979). For example, the disclosure of traumatic or stressful events has been shown to reduce stress and lead to positive health outcomes in those with diabetes (Bodor, 2002) and breast cancer (Stanton et al., 2002), among others. Inhibiting repressive thoughts or emotions, rather, may be detrimental to both physical and psychological health (H. S. Goldstein, Edelberg, Meier, & Davis, 1988; Gross & Levenson, 1997; Larson & Chastain, 1990). While some studies suggest that emotional expression in the form of "truth telling" may cause psychological harm to individuals (Brounéus, 2010), the literature presents a plethora of evidence confirming the negative effects of a lack of emotional expression, such as social concerns, overall psychological dysfunction, and lack of value-congruent behaviors (Frankl, 1959; Pennebaker, 1989; Pennebaker & Beall, 1986; Schulenberg, Hutzell, Nassif, & Rogina, 2008; Wilson & DuFrene, 2009). These resulting negative outcomes may lead to detrimental effects on health (Pennebaker & Beall, 1986). Individuals having experienced a traumatic or stressful life event are significantly more likely to repress thoughts and feelings about their experience compared to individuals who have not experienced such events, thereby subjecting them to potential negative outcomes caused by a lack of emotional expression (Bodor, 2002).

Expressive Writing as Effective Emotional Expression

Pennebaker and Beall (1986) first showed that emotional expression can be both experimentally manipulated and have positive benefits to participants. In their seminal study, they randomly assigned participants to several writing groups, including writing about

a "stressful or traumatic" life event or a neutral event. As such, the content of the writing likely varies widely based on the contextual factors (e.g. topic, setting, sample, health concern). The group that disclosed both regarding their trauma and the emotions surrounding said trauma later showed a reduction in health visits. Pennebaker has replicated the use of expressive writing across a number of studies ranging from improved health (Pennebaker, Colder, & Sharp, 1990; Pennebaker, Kiecolt-Glaser, & Glaser, 1988) to improvements in school (Pennebaker & Francis, 1996) and work (Francis & Pennebaker, 1992). Others have expanded this work to show positive effects on mood (Schoutrop, Lange, Hanewald, Davidovich, & Salomon, 2002) and asthma (Smyth, Stone, Hurewitz, & Kaell, 1999); however, several controlled studies have shown to not replicate (Harris, Thoresen, Humphreys, & Faul, 2005) or null effects (Gidron, Peri, Connolly, & Shalev, 1996; Walker, Nail, & Croyle, 1999).

The idea that a brief, controlled writing task can have numerous positive health and psychological benefits can certainly be controversial, given the existing literature. For example, Henry, Schlegel, Talley, Molix, and Bettencourt (2010) found that expressive writing only benefited a rural population for those individuals surviving breast cancer on physical and psychological health outcomes, while Lancaster, Klein, and Heifner (2015) found no significant evidence that expressive writing can be considered an effective approach in measuring posttraumatic growth. Additionally and as mentioned, Brounéus (2010) found that "truth telling" caused harm to individuals in a forensic setting. Regardless, the concept remains interesting due to the nature and inexpensive implementation of expressive writing. Many individuals who have experienced traumatic events do not wish to disclose their feelings regarding the events with others. Additionally, those who do not meet diagnostic criteria (e.g. subclinical symptoms) are sometimes neglected despite probable suffering (Wilson & DuFrene, 2009). However, by utilizing expressive writing as a personal method of treatment, individuals are able to effectively express their emotions while avoiding talking to another individual or clinician about the traumatic event (Smyth, 1998). Pennebaker (1993)

found that experimental conditions assigned to participate in an expressive writing task generally report more positive changes than those in control conditions. Some controversy has been observed over whether or not writing about a formerly disclosed event is more effective than writing about an undisclosed event. M. A. Greenberg and Stone (1992) conducted an experiment where they separated participants into three groups: writing about a formerly disclosed trauma, writing about an undisclosed trauma, and a control group. They found no difference between groups in efficacy. However, they did find that those who disclosed more severe traumas reportedly experienced fewer physical symptoms at follow up, which suggests that the type of trauma revealed can play a significant impact on symptom reduction and physical health. A review of current meta-analyses is presented in a subsequent section.

Possible Mechanisms Underlying Expressive Writing Efficacy

In order to understand why expressive writing is considered to be efficacious, one must examine the cognitive, social, and behavioral processes by which it allows information processing. Pennebaker et al. (1990) discovered that individuals who benefited from expressive writing attributed their success from the writing task to a renewed sense of understanding. Further, Pennebaker (1993) conducted a textual analysis on expressive writing content and found that those who were more successful during the task used causation words. Pennebaker considered expressive writing as a way for individuals to effectively process the event in their minds, which may explain the aforementioned renewed sense of understand and excess of causation-oriented words. Aside from theories related to cognitive-processing and inhibition, there are a number of other theories related to emotional disclosure that warrant explanation. The first is the social integration model (Pennebaker & Graybeal, 2001). This model suggests emotional disclosure can have a positive impact on how people interact in their environment. This increased environmental interaction has been shown to have positive benefits on health (Frattaroli, 2006). Second, expressive writing

101 parallels exposure therapy for phobias and posttraumatic stress disorder, which suggests that

102 repeatedly exposing oneself to the fear or trauma can reduce the negative emotions

103 associated with that fear or trauma (Meshberg-Cohen, Svikis, & McMahon, 2014). Given

104 that exposure therapy has been shown to be effective for reducing symptoms of

105 posttraumatic stress (PTS; Sloan, Marx, & Epstein, 2005), one would expect individuals in

106 these studies to experience a reduction in PTS symptoms after taking part in an expressive

107 writing task. Third, Wilson and DuFrene (2009) discussed how the nonjudgmental

108 acceptance of emotions leads to positive health benefits by promoting value-congruent

109 behavior, one of the main facets of Functional Contextualism theory and Logotherapy

110 (Frankl, 1959; Schulenberg et al., 2008). Indeed, emotional expression in the form of

111 expressive writing could be considered a form of nonjudgmental acceptance, although it may

112 not necessarily lead to behavior change. Engaging in value-congruent behavior in the

113 presence of inevitable emotional experiences is also the most crucial tenet in Logotherapy.

114 Finally, a recently proposed theory that may help explain positive outcomes is referred to as

115 a distance percpetive (Kross and Ayduk (2011)). This theory posits that, when individuals

116 adopt a psychologically distanced perspective, they are better able to better understand

117 their life situation. In sum, it seems likely that there are multiple underlying mechanisms

118 that account for the beneficial outcomes associated with expressive writing described below.

119 Indeed, the wide range of theroetical perspectives provide further evidence which suggests

120 that expressive writing is applicable in a variety of contexts. Previously conducted

121 meta-analyses, however, present varying results.

Meta-Analytic Techniques

123 Meta-analyses allow researchers the opportunity to collectively examine the efficacy of

124 different psychological interventions/tasks on outcome variables (Borenstein, Hedges, &

125 Rothstein, 2007; Glass, 1976; Hedges, 1982). Although many studies produced positive

126 outcomes associated with expressive writing, some of these studies tend to neglect important

127 questions, the most important of which is whether or not the effect sizes are meaningful

128 (Smyth, 1998). Meta-analyses are a technique that allows researchers to pool studies to

129 examine an overall, weighted, population effect (Borenstein et al., 2007). Several

130 meta-analyses of expressive writing and emotional expression have been explored that

131 warrant considerable explanation: Smyth (1998), Frisina, Borod, and Lepore (2004),

132 Frattaroli (2006), Reinhold, Bürkner, and Holling (2018), Van Emmerik, Reijntjes, and

133 Kamphuis (2013) and Mogk, Otte, Reinhold-Hurley, and Kröner-Herwig (2006). These

134 meta-analyses have laid a foundation for exploring the effects of writing on psychological

135 outcomes. The meta-analyses described sequentially below focused on experimental versus

136 control group effect sizes or p-values, rather than emphasizing change for an writing group.

137 This focus is likely because of the analyses provided in these publications, especially when

138 using randomized controlled trial research designs. While this design is the gold standard for

139 medicine, the current meta-analysis sought to examine the magnitude of change for

140 participants who experienced an expressive writing task. For example, a comparison group

141 may increase their quality of life scores by two points in a controlled study, while the

142 experimental group increases their quality of life scores by four points; thus, creating a

143 significant difference in change between the two groups. This information is valuable, but it

144 does not tell the reader the magnitude of the change for the writing group, wherein four

145 points might only be a small effect when examined within the group who received the

146 writing task. This study will focus on changes across time for groups who received the

147 expressive writing task to determine what size of effects one might expect given a specific

148 measurement schedule (i.e., one to three months, three months to six months, etc.).

149 Additionally, the current meta-analysis will examine the change in effects across

150 measurement time to determine whether or not effects decrease over time.

151   Smyth (1998) conducted the seminal meta-analysis regarding the efficacy of expressive

152 writing. They included studies utilizing an expressive writing group and control group

153 (neutral topic). This particular analysis examined the efficacy of expressive writing on

154 psychological well-being, general health, and physical functioning. In sum, 13 studies/effect

155 sizes were included, and the authors found an overall medium effect size, $d = 0.47$, for the

156 experimental group compared to the control group. A later meta-analysis conducted by

157 Frisina et al. (2004) expanded these analyses. They only included studies utilizing clinical

158 samples and employing the paradigm adapted by Pennebaker and Beall (1986). This

159 meta-analysis included 9 studies in total and found an effect size of $d = .19$ for

160 health-related outcomes and $d = .07$ for psychological outcomes. The next meta-analysis

161 regarding expressive writing was conducted by Mogk et al. (2006) and aimed to update the

162 state of the literature on expressive writing. Similar to previously-conducted analysis, they

163 included studies employing Pennebaker's paradigm on experimental and control groups.

164 Additionally, they only included studies with a 4-week follow up that included at least 10

165 participants. In sum, 30 studies met their criteria. They found nonsignificant effects on

166 somatic and psychological health outcomes and concluded that expressive wrting does not

167 benefit health. These findings corroboate those from Frisina et al. (2004). Frattaroli (2006)

168 conducted perhaps the most notable meta-analysis to data examining the efficacy of

169 emotional disclosure on the folliwing constructs using only randomized and control

170 conditions: psychological health, physiological functioning, reported health, health behaviors,

171 and general functioning/life outcomes. Additionally, their meta-analysis employed random

172 effects models. Prior meta-analysis employed fixed effects models, which assume that all

173 studies assess the same "true" population effect size, which may be an untenable assumption

174 across different assessment and populations (Borenstein et al., 2007). Random effects models,

175 rather, estimate the mean of a proposed distribution of population effect sizes, which may

176 vary by subject type or research design. They included a wide range of studies, $N = 146$.

177 Individual studies were again collapsed into one publication effect size, although these effects

178 were also examined separately by health outcome. Overall, Frattaroli (2006) found a

179 weighted r effect size of .08 for all outcomes combined, which would be considered small.

180 Additionally, they examined potential moderators and found larger effect sizes for the

following samples: those with physical health problems, those with a history of having

experienced traumatic or stressful events, samples not including college students, samples

where expressive writing tasks were conducted at home and in private settings, paid

participants, more male participants, and fewer participants (see Frattaroli, 2006 for a

complete list of moderators). A recent analysis conducted by Van Emmerik et al. (2013)

employing Pennebaker's paradigm found included 6 eligible studies that compared treatment

to control groups. In regards to inclusion criteria, they included studies where participants

had a diagnosis of Acute Stress Disorder (ASD) or Posttraumatic Stress Disorder (PTSD).

They found that those who participated in the expressive writing group experienced

short-term reductions in PTS and depressive symptoms. combined Hedges' $g = .81$. The

most recently published meta-analysis was conducted by Reinhold et al. (2018) and

examined the effects of expressive writing on depression by randomizing participants to

conditions (expressive writing vs. control). They included 39 randomized controlled trials

and excluded individuals with diagnoses of PTSD. This study did not support utilizing

expressive writing for depression outcome measures. Further, they found that expressive

writing did not yield any type of long-term effect on depression outcomes.

Posttraumatic Stress

Posttraumatic Stress Disorder (PTSD) is a disorder involving re-experiencing thoughts

or experiences after a traumatic event or experience. This generates a context where

individuals are prone to affect-related deficiencies and maladaptive behaviors (American

Psychiatric Association, 2013). DSM-5 criteria are based on 20 symptoms structured into

four different subsets in those having experienced a traumatic event. These subsets are as

follows: re-experiencing, avoidance, negative alterations in cognition and mood, and

increased arousal (Crespo & Gomez, 2016). While the renewed DSM-5 criteria are commonly

employed, the current meta-analysis considers studies using DSM-IV criteria. DSM-IV

criteria are similar and include the following: exposure to a traumatic event, re-experiencing

207 (intrusion), avoidance, and increased arousal (American Psychiatric Association, 2013).

208 Further, the studies employed in the current meta-analysis are divided according to these

209 subsets (arousal, intrusion, and avoidance). PTSD affects a wide variety of groups, a few of

210 which are sexual assault survivors (Klump, 2008), Iraq and Afghanistan war veterans (Gentes

211 et al., 2014), and those exposed to natural disasters (Wang et al., 2000). Research conducted

212 on the efficacy of expressive writing on PTSD symptoms has been less successful and shows

213 outcomes that are not as effective as other studies. While PTSD symptoms decreased after a

214 written emotional disclosure task, this decrease was not significantly different than a control

215 group change (Sloan, Marx, & Greenberg, 2011). However, other studies show contradicting

216 results. For example, Sloan, Marx, Epstein, and Lexington (2007) examined individuals with

217 at least moderate PTSD symptom severity and found that individuals assigned to an

218 emotional expression writing condition reported fewer PTSD and depression symptoms

219 during follow up. Additionally, in 2012, Sloan, Marx, Bovin, Feinstein, and Gallagher (2012)

220 adapted their writing protocol to focus primarily on the emotions, meaning, and "hot spots"

221 associated with the trauma. They referred to this as the written exposure therapy (WET)

222 protocol, distinguishable from that adapted by Pennebaker and Beall (1986).

223 Di Blasio et al. (2015) recruited women who had just given birth and assessed them a

224 few days after experiencing childbirth along with a three-month follow-up. Results showed

225 that women who had participated in the expressive writing task had lower depression and

226 posttraumatic stress symptoms than the group assigned to a neutral writing condition.

227 Additionally, regression models showed that expressive writing was significantly linked to a

228 reduction of PTSD symptoms across different dimensional levels of symptom severity.

229 Contradicting the work conducted by Sloan et al. (2011), these results suggest that

230 expressive writing may be an effective tool for women managing postpartum distress.

231 However, it is important to note that only 20 of the 113 women recruited for this study

232 qualified for a diagnosis of PTSD. This limitation suggests that those with moderate distress

233 could perhaps benefit more from an expressive writing task than those diagnosed with or

234 meeting the qualifications for PTSD. It may also explain the differences in results between

235 the two studies, as Sloan et al. (2011) found that those with a clinical diagnosis of PTSD did

236 not respond to an emotional disclosure writing task. Di Blasio et al. (2015) found that those

237 who reported mild symptomology better responded to the task than those meeting the

238 criteria for a diagnosis of PTSD. Perhaps it is more advantageous to examine effect sizes

239 separately in with diagnoses of PTSD and subclinical symptoms.

240 **Posttraumatic Growth**

241   While the literature mostly discusses potentially harmful outcomes to traumatic events

242 such as emotional distress, traumatic events also provide opportunities for personal growth

243 (Aslam & Kamal, 2013). Traumatic events, either natural or human-inflicted, can lead to

244 positive outcomes by allowing the individual to take a different perspective (Cobb, Tedeschi,

245 Calhoun, & Cann, 2006; Taku, Calhoun, Cann, & Tedeschi, 2008). The relationship between

246 positive growth after a traumatic event and symptom reduction is unclear, as it is a complex

247 process. Thus, it is necessary to examine how expressive writing might influence each

248 variable separately, which is one of the key goals of this meta-analysis (Slavin-Spenny, Cohen,

249 Oberleitner, & Lumley, 2011). Models receiving empirical support within the last decade

250 suggest that traumatic events offer opportunities for both negative and positive experiences

251 (Tedeschi & Calhoun, 1995; Weiss, 2002). Posttraumatic Growth (PTG) is a positive

252 experience after a traumatic event (Aslam & Kamal, 2013; Yilmaz & Zara, 2016).

253 Specifically, PTG is classified as broad cognitive benefits that are seen after a traumatic

254 experience. These benefits can be categorized into building closer relationships, examining

255 new possibilities, appreciating life, recognizing personal strengths, and undergoing spiritual

256 changes. (Dursun, Steger, Bentele, & Schulenberg, 2016; Tedeschi & Calhoun, 2004).

257   PTG is associated with a variety of desired outcomes (Dursun et al., 2016). PTG has

258 been studied in those experiencing natural disasters, war, and other harms such as sexual

259 assault. Finally, PTG has been studied in those experiencing medical diagnoses such as

different types of cancer and diseases. Although the relationship between PTG and symptom reduction is not yet fully understood, perhaps expressive writing allows the individual to fully comprehend the event. Pennebaker and Graybeal (2001) speculated that expressive writing allows an individual to feel more connected with his or her surroundings. Although this speculation does not directly explain positive outcomes after an expressive writing task, perhaps individuals gain a better appreciation for life after gaining a better sense of connectedness with that individual's surroundings.

Quality of Life

Quality of Life (QOL), according to Theofilou (2013) is an evaluation of the "goodness" that an individual experiences, separated into domains of reactions to life events, disposition, life fulfillment, and satisfaction with life experiences. More generally, QOL refers to an individual's attitude towards the target life situation (Costanza et al., 2007), delineated into objective and subjective components. Objectively, QOL refers to components outside of an individual and measurable by others, while subjective QOL is an individual's assessment of his or her own experiences (Costanza et al., 2007). The current meta-analysis will focus solely on the subjective components of QOL, as it is obtainable through questionnaires. Pennebaker and Graybeal (2001) suggested that expressive writing allows one to feel more connected with their surroundings. Further, they explain that expressive writing allows people to see things in a different way and better understand themselves. By understanding a traumatic or stressful event, one is said to see things differently and perhaps look at the situation with a more positive mindset. The changes that occur after expressive writing may also allow one to find meaning in the traumatic event, thereby increasing the QOL of that individual (Frankl, 1959). Higher QOL may be considered a type of PTG, which is why the current meta-analysis sought to examine the efficacy of studies utilizing expressive writing to improve QOL and PTG in the same study.

Current Meta-Analysis

The purpose of the current meta-analysis is to examine studies employing expressive writing procedures using Pennebaker's paradigm on variables relevant to the field of positive psychology (PTG and QOL) and PTS, with effect sizes separated by those having and not having a diagnosis of PTSD. Based on recently published literature regarding efficacy of expressive writing for different levels of PTSD symptoms, this is an important facet to consider, (Reinhold et al., 2018). Additionally, no meta-analysis has been conducted that examines the efficacy of expressive writing on positive outcome variables such as PTG and QOL. The meta-analyses described sequentially above focused on experimental versus control group effect sizes or p-values, rather than emphasizing change for the expressive writing group. This focus is likely because of the analyses provided in these publications, especially when using randomized controlled trial research designs. While this design is the gold standard for medicine, the current meta-analysis sought to examine the magnitude of change for participants who experienced an expressive writing task. For example, a comparison group may increase their quality of life scores by two points in a controlled study, while the experimental group increases their quality of life scores by four points; thus, creating a significant difference in change between the two groups. This information is valuable, but it does not tell the reader the magnitude of the change for the writing group, wherein four points might only be a small effect when examined within the group who received the writing task. This study will focus on changes across time for groups who received the expressive writing task to determine what size of effects one might expect given a specific measurement schedule (i.e., one to three months, three months to six months, etc.). This analysis should present researchers with a renewed examination of the efficacy of expressive writing on the aforementioned variables using newer meta-analytic techniques. Newer methods of meta-analysis, including p-curve (Simonsohn, Nelson, & Simmons, 2014; Simonsohn, Simmons, & Nelson, 2015), p-uniform (Aert, Wicherts, & Van Assen, 2016), PET-PEESE (Stanley & Doucouliagos, 2014), selection models (Vevea & Hedges, 1995), and

trim and fill methods (Carter & McCullough, 2014) allow for better estimation of meta-analytic effect sizes. These analyses would be best performed by examining each potential effect separately, rather than averaging effects of each publication into one study effect size (a common trend in the previously mentioned meta-analysis). In addition to an estimate of overall effect sizes using updated techniques, the current meta-analysis estimates power for effects on writing groups, as research has shown a consistent underpowering of psychological studies, combined with a misunderstanding of the sample size needed for adequately powering one's work (Bakker, Hartgerink, Wicherts, & Maas, 2016).

Method

Data Collection

Studies were collected through online databases, such as PsycINFO and Google Scholar, using the following search terms and their combinations: *Posttraumatic Growth, PTG, Quality of Life, QOL, Posttraumatic Stress, PTS, Expressive Writing, Emotional Disclosure, Narrative Writing.* Within these articles, the change in outcome variables (PTS, PTG, QOL) from pre- to post-test was the dependent variable of interest. Generally, groups were separated into an experimental and control group and then examined at different time points. For purposes of this meta-analysis, only participants assigned to the experimental condition were examined due to having received the expressive writing task. If a study included multiple assessment time points, then these measurements were examined sequentially (i.e., time 1 to time 2, time 2 to time 3) to determine change across time for the dependent variable.

220 citations focusing on PTS, PTG, and QOL were identified through the literature search and previous meta-analyses. After screening these studies, forty-five articles were retained for containing the appropriate information for this meta-analysis. A complete list of excluded articles can be found at https://osf.io/4mjqt, along with reasons why they were excluded. Generally, studies were included if they utilized the expressive writing paradigm

338 adapted by Pennebaker and Beall (1986), included relevant numbers to compute an effect

339 size, and included the relevant outcome variables. After having two reviewers independent

340 code articles, 202 effect sizes were calculated from the forty-five studies. On average, each

341 study represented $M = 4.49$ ($SD = 3.50$) effects, ranging from 1 to 16 effects. 144 effects

342 were calculated for PTS, 21 for PTG, and 37 for QOL.

343 **Calculations for Effect Size, Variance, and Confidence Intervals**

344   For our purposes, we used Cohen's (1988) standards for nomenclature for small (0.20),

345 medium (0.50), and large (0.80) d values, although it is important to note that Cohen

346 himself suggested that these values should be based on the area of study. Generally, however,

347 these effect size criteria are used within the social sciences. Each study implemented a

348 pre-test to post-test style repeated measures design, usually with paired t-tests, ANOVA, or

349 regression analyses. The means, standard deviations, and N values were collected from each

350 study. In general, Cohen's d values were calculated using the following formula for paired t

351 using means and standard deviations:

$$d_{av} = \frac{M_1 - M_2}{\frac{SD_1 + SD_2}{2}}$$

352   This equation is described in detail in Cumming (2012) as an alternative to the

353 traditional calculation of d for paired samples t, wherein the denominator is the standard

354 deviation of the difference scores:

$$d_z = \frac{M_1 - M_2}{SD_{diff}}$$

355   This equation for d_{av} not only allows for calculations from published articles that do

356 not include SD_{diff} (i.e., most articles included), but also has been shown to be less upwardly

357 biased than d_z. Alternative formulas include controlling for r between paired levels, as

358 described in Lakens (2013); however, these values were not available in the selected articles,

359 and Lakens also recommends d_{av} as an effect size for paired designs. When only mean

360 differences and standard deviation of the difference scores were available, the second

361 equation for d_z was used.

362   We planned to use traditional and newer methods of meta-analysis, following guidelines

363 from Cooper, Hedges, and Valentine (2009) and Borenstein et al. (2007), as well as Aert et

364 al. (2016). Sampling variance of the effect sizes were estimated using the *escalc()* function

365 from the *metafor* package in *R* (Viechtbauer, 2010). The variance formula was originally

366 published in S. B. Morris and DeShon (2002) and is shown below:

$$v = \frac{1}{n}\left(\frac{n-1}{n-3}\right)(1 + n * d^2) - \frac{d^2}{[c(n-1)]^2}$$

367   In this formula, n is the number of paired observations, d is the calculated effect size,

368 and c is a correction factor, wherein *df* are $n-1$ (Hedges, 1982):

$$c = 1 - \frac{3}{4 * df - 1}$$

369   We used the *metagen()* function in the *metafor* package to calculate both fixed and

370 random effects models, which uses standard error of the effect to calculate overall estimates

371 of an effect and their confidence intervals. Thus, we took the square root of the variance

372 estimate for standard error. Given these calculations, the goal of this analysis was to

373 calculate a combined effect size, along with a confidence interval for study planning and an

374 assessment of the literature. A fixed effects model requires the assumption that there is a

375 true population effect size across all studies. By including multiple measures of psychological

376 outcomes, this assumption may be tenuous, and therefore, a random effects model was also

377 calculated. In random effects models, the true effect is assumed to vary across studies

378 (Borenstein et al., 2007). For a fixed effects model, the effect sizes are weighted by their

379 inverse variance (v; Sánchez-Meca & Marín-Martínez, 2008), which is calculated

380 automatically in *metafor* by:

$$w_i^{FE} = \frac{1}{v}$$

381   The advantage to this procedure is that analyses are weighted by their precision, that

382 is, that studies with more information (often, larger samples), are given larger weights in the

383 overall estimated effect size (Borenstein et al., 2007). Random effects models are also

384 weighted by inverse variance, with an additional correction for variance between studies, τ^2_{DL},

385 as described by DerSimonian and Laird (1986):

$$w_i^{RE} = \frac{1}{v + \tau^2_{DL}}$$

386   Confidence intervals were calculated in two ways for this study. Cumming (2012),

387 Kelley (2007), and Smithson (2001) have shown that the distribution of d values are

388 non-normal, and thus, CIs should be estimated using the non-centrality parameter and a

389 non-normal distribution. These values were calculated using the functions in the *MOTE*

390 library which iteratively estimates the appropriate non-centrality parameter and converts

391 back to d values (i.e., non-centrality parameter divided by the square root of n; Buchanan,

392 Valentine, & Scofield, 2017; Smithson, 2001, 2003). However, the *metafor* package in R uses

393 central distributions to estimate CIs for each study and overall effect sizes. Therefore, we

394 present both sets of values for the interested reader, as meta-analytic procedures have not

395 implemented non-central distributions of effect sizes.

396 **Additional Meta-Analytic Techniques**

397   **p-Curve and p-Uniform.**   We used *p*-curve.com to conduct a *p*-curve analysis

398 (Simonsohn et al., 2014). The purpose of this type of analysis is to detect true effects.

399 Specifically, *p*-curve is used to reveal possible *p*-hacking in published literature in order to

400 decipher whether or not a true effect exists. Broadly, *p*-hacking occurs when researchers use

401 questionable research practices to create significant results by manipulating dependent

402 variables or covariates. Additionally, authors may add participants if the initial findings are

403 not significant (Bruns & Ioannidis, 2016). Researchers may also decide to exclude

404 participants for final analyses if that exclusion leads to a significant difference (L. K. John,

Loewenstein, & Prelec, 2012). Thus, it is necessary to distinguish between true and false effects in order to effectively interpret effect sizes corresponding to those p-values. p-curve accomplishes this task by examining the distributions of the published p-values. If an effect exists, or rather the results should be interpreted as presented, the distribution of p-values will be positively skewed (Simonsohn et al., 2014). If, however, no effect exists, then the distribution of p-values will be flat. p-curve analyses ultimately provide evidence of p-hacking in groups of studies and has become an important tool for interpreting meta-analyses. In order to accurately estimate effect sizes because of scrutiny associated with effect size estimation of p-curve, we also conducted p-uniform. p-uniform analyses, too, are interpreted by examining the distribution of p-values in a set of studies (Aert et al., 2016). However, it is assumed that the population effect size equals the effect size from the dataset. Because of this assumption, the population effect size is referred to as uniform. This analysis also examines for publication bias and presents the researcher with a corrected effect size. Publication bias occurs when only select studies are published, usually only significant studies, although many factors can bias a study's publication (McShane, Böckenholt, & Hansen, 2016). p-uniform was calculated from code provided by Van Aert (2017) on GitHub.

PET-PEESE.   Originally, meta-analyses relied on the calculation of Egger's regression test which examined the relationship of the standard error (predictor) to the effect size estimates (criterion). In this regression, the intercept values were used to determine if effect size measures were different than zero, by providing a meta-analytic estimate (Egger, Davey Smith, Schneider, & Minder, 1997; Stanley, 2005). PET-PEESE analyses examine for publication bias by adapting parts from Egger's traditional regression tests: PET (Precision Effect Test) and PEESE (Precision Effect Estimate with Standard Error, Carter & McCullough, 2014). PET is a more reliable test of publication bias with effect size estimates of zero, $b_0 = 0$, while PEESE is more accurate with non-zero effect size estimates, $b_0 \neq 0$ (Stanley & Doucouliagos, 2014). PET-PEESE was calculated using Hilgard's (2016) code provided on GitHub.

432   **Selection Models.**   Selection model analyses provide the researcher with a test of

433 publication bias and effect size estimates using maximum likelihood estimation (Vevea &

434 Hedges, 1995; Vevea & Woods, 2005). Using selection models, researchers are able to

435 discover effect size estimates as well as evidence of publication bias (McShane et al., 2016) by

436 using a mixed general linear model to estimate these values. Selection models were

437 calculated with the *weightr* package in *R* (Coburn & Vevea, 2017).

438   **Trim and Fill.**   Trim and Fill analyses, in contrast to PET-PEESE, regress standard

439 error (criterion) and effect size estimates (predictor). Specifically, the purpose of Trim and

440 Fill techniques is to examine whether or not publication bias may influence the regression

441 equation (Carter & McCullough, 2014). Effect sizes and standard error terms are graphically

442 displayed on x and y-axes, respectively, in a funnel plot. If this graphical representation

443 indicates asymmetry, considered a gap of missing data points in the lower center area of the

444 plot, the study set can be assumed to have studies that are both non-significant and small in

445 sample size (Van Assen, Van Aert, & Wicherts, 2015). This funnel is then trimmed until

446 symmetry is achieved. Missing studies from the symmetrical graph are imputed (filled) while

447 maintaining the given symmetry (Duval & Tweedie, 2000). The meta-analytic effect size is

448 then estimated from the trimmed and filled funnel plot. Trim and fill analyses, as well as

449 funnel plots included below, were calculated with the *metafor* package.

450   **Results**

451 **PTS**

452   **Overall Effect Size.**   As described above, both fixed effects and random effects

453 models with centralized confidence intervals are presented in Table 1. Studies were examined

454 for potential outliers using the *metafor* package in *R*. This package calculates traditional

455 regression influence values, such as Cook's and hat values (J. Cohen, 1988). These values

456 indicate change in overall meta-analytic model with and without the effect; thus,

457 determining their impact on the pooled effect size (Viechtbauer, 2010). Because published

458   studies likely represent the range of the sampling distribution of effect sizes, we included the

459   analyses with and without outliers to present evidence for both paths a researcher might

460   take when examining an overall effect.

461   Three outliers were detected with this procedure, all showing very large effect sizes,

462   average $d = 1.63$. The fixed and random effects estimates without these points are also

463   included in Table 1. Figures 1, 2, 3, and 4 portray the effect sizes for PTS studies, separated

464   by intrusions, avoidance, hyperarousal, and total scores for easier viewing (i.e., over 100+

465   effect sizes did not fit easily on one combined graph). Although these categories are not

466   reflective of updated DSM-5 criteria, researchers have not yet conducted enough studies using

467   expressive writing on PTS with updated PTSD criteria to warrant a meta-analysis. Name

468   acronym coding can be found in the data online. This forest plot includes the non-centralized

469   confidence interval calculated from the *MOTE* library (Buchanan et al., 2017). Shape size

470   indicates study weight, and these values were taken from the overall random effects

471   meta-analysis and normalized by dividing by the mean weight. The dashed lines indicate the

472   average non-weighted lower and upper confidence interval limit for the non-centralized

473   estimates. Overall, PTS studies include a small effect size that appears to be significantly

474   greater than zero across all estimate types (fixed, random, with or without outliers).

475   **Homogeneity.**   A prerequisite for newer meta-analytic techniques includes the

476   assessment of homogeneity of the effects (Aert et al., 2016). Using the *metafor* package in *R*,

477   we calculated the Q-statistic and the I^2 index (Cochran, 1954; Huedo-Medina,

478   Sánchez-Meca, Marín-Martínez, & Botella, 2006). Significant values imply inconsistencies

479   across the variable or variables of interest and are represented by Q. In contrast, I^2 indicates

480   the percentage of heterogeneity along with a 95% CI. Both can, however, be biased with a

481   small number of experiments included for analyses (Higgins, Thompson, Deeks, & Altman,

482   2003; Huedo-Medina et al., 2006). Thus, we sought to calculate an overall level of

483   heterogeneity after examining each variable separately before and after excluding outliers.

484   For PTS studies including outliers, we found significant heterogeneity, $Q(143) = 639.98$, $p <$

.001 and $I^2 = 77.7$, 95% CI[73.9 - 80.9]. These values were reduced slightly with the exclusion of outliers, $Q(140) = 519.75$, $p < .001$ and $I^2 = 73.1$, 95% CI[68.2 - 77.2].

Power. Power was calculated in two different ways using the *pwr* package in *R* (Champely, 2016). *Post hoc* power was first calculated using sample size and effect size statistics from each individual study. Additionally, we calculated power using the study sample size and estimated overall effect size from the random effects model with and without outliers, as explained by G. Francis (2012) and G. Francis (2014). The first estimate indicates the likelihood of finding an effect from our sample statistics, while the second indicates the likelihood of finding the true population effect size. If each study had been conducted on only the change in the experimental group, 45.1% of studies would have been considered significant at $\alpha < .05$. The average power of these studies based on their original study characteristics was .46 ($SD = .36$). Power for the random-effects meta-analytic effect size with outliers was .47 ($SD = .24$) and without outliers was .42 ($SD = .23$). Therefore, power consistently was around 40-50% for studies examining PTS, regardless of outlier effects. In these studies, only 26.4% achieved recommended 80% power for their found effect size, a smaller 16.7% for the random-effect outlier effect size, and even smaller 6.9% for power calculations on the random-effect size without the outliers.

Other Meta-Analytic Estimates. As noted in Aert et al. (2016), *p*-curve and *p*-uniform analyses are upwardly biased when heterogeneity is high. Therefore, we use caution when interpreting these analyses on PTS outcomes. As seen in Table 1, the estimates for *p*-uniform were higher than other techniques, likely because of the focus on significant *p*-values and the great degree of heterogeneity described earlier. *P*-curve pictures can be found at https://osf.io/4mjqt/ online, and this analysis indicated evidentiary value at $p < .001$. Additionally, the *p*-uniform analysis indicated that there was likely no publication bias present, $Z = -5.02$, $p = 1.000$. When examining the PET analysis, we found that the intercept was significant, which indicated that PEESE was likely a better estimator of the meta-analytic effect size. PEESE estimates were lower than the original meta-analytic

estimate, but confidence intervals indicated that the effect is small to medium, and still

larger than zero. Selection models indicated a larger effect size, especially with the

random-effects models, and these effects were influenced by the outliers found in the

published studies. Trim and fill models are shown in Table 1, and figures are included online.

Nineteen missing studies were imputed for both models with and without outliers. Across all

these effect size estimates, we found that expressive writing was likely to decrease PTS

symptoms in a small to moderate way. The correlation of effect size with time between

measurement times was $r = -.16$, 95% CI $[-.32, .00]$, $t(142) = -1.99$, $p = .049$, and

$r = -.15$, 95% CI $[-.30, .02]$, $t(139) = -1.75$, $p = .082$ without outliers. This result

indicated that the effect of expressive writing slightly decreased across time.

PTG

Overall Effect Size.   Both fixed and random effects models with centralized

confidence intervals for PTG are presented in Table 2. When examining expressive writing

on PTG, no outliers were detected. Fixed and random effects estimates are included in Table

2, while Figure 5 shows effect sizes for PTG studies where shape size indicates the

normalized weight of the study. Dashed lines indicate non-weighted lower and upper

confidence intervals for non-centralized estimates. Overall, PTG studies indicated a

negligible to small effect size across both random and fixed effects models, and the

non-centralized confidence intervals indicated an effect that crossed zero.

Homogeneity.   Using the *metafor* package in R, we calculated both a Q statistic and

I^2 index. Since PTG studied did not contain any outliers, we did not calculate two separate

analyses to examine heterogeneity both with and without outliers. We did not find significant

heterogeneity across PTG studies, $Q(20) = 14.18$, $p = .82$ and $I^2 = 0.0$, 95% CI$[0.0 - 25.3]$.

Power.   First, we calculated *post hoc* power using both sample and effect size

statistics from individual studies. Individual studies examining change in experimental

groups showed that 9.5% of studies would have been considered significant at $\alpha < .05$.

Average power of PTG studies was .15 ($SD = .16$). 0.0% achieved recommended 80% power for their found effect size. Additionally, we calculated power using study sample size and estimated effect size from our random effects model. Power for the true effect size was .08 ($SD = .02$). Again, 0.0% achieved recommended 80% power.

Other Meta-Analytic Estimates. Due to no heterogeneity across PTG studies, we can use both *p*-curve and *p*-uniform analyses with more confidence. A pictorial representation of *p*-curve can be found at https://osf.io/4mjqt/. This analysis did not indicate evidentiary value, $p = .75$, as only two of the results would be considered significant at $\alpha < .05$. *p*-uniform estimates are presented in Table 2. Specifically, these analyses indicated that there was no publication bias present, $Z = 0.70$, $p = .243$. The *p*-uniform estimates of the effect size for PTG were negative, in contrast to the fixed and random effects overall model. The confidence interval for this analysis indicates a wide range of possible effects. In examining PET-PEESE analyses, we did not find a significant intercept, indicating that PET is most likely a better effect size estimator. PET analyses indicated that the effect size is negligible to small, with our confidence interval crossing zero. These results corroborated our original effect size calculations. Selection models indicated negligible to small effect sizes, again wherein the confidence interval includes zero effect. Trim and fill models are shown in Table 2, and figures are included online. Zero studies were imputed for our model, and thus, the effect size estimate is the same as the overall model. Across techniques, we found that expressive writing has little to no effect on PTG. The correlation of effect size across measurement times in PTG studies at subsequent time points was $r = .09$, 95% CI $[-.36, .50]$, $t(19) = 0.38$, $p = .707$, and no change over time was found.

QOL

Overall Effect Size. Finally, for QOL, both fixed and random effects models with centralized confidence intervals are presented in Table 3. Two outliers were detected with this procedure, average $d = $ -0.07. While the average effect of these outliers indicates a small

564 number, it is important to note that these two outliers were the largest positive and negative

565 effects found from the Possemato, Ouimette, and Geller (2010) study. Fixed and random

566 effects estimates without these points are also included in Table 3, while Figure 6 shows

567 effect sizes for QOL studies. Overall, QOL studies indicated a negligible to small effect that

568 showed a non-significant decrease in quality of life as a result of expressive writing.

569 **Homogeneity.**   For QOL studies including outliers, we found significant

570 heterogeneity from our random effects model, $Q(36) = 200.09$, $p < .001$ and $I^2 = 82.0$, 95%

571 CI[75.9 - 86.5]. After excluding outliers, our random effects model still indicated

572 heterogeneity, $Q(34) = 93.18$, $p < .001$ and $I^2 = 63.5$, 95% CI[47.6 - 74.6].

573 **Power.**   In conducting *post hoc* power using sample and effect size statistics from

574 individual studies, we found that 21.6% of studies would have been considered significant at

575 $\alpha < .05$. Average power based on actual study characteristics was .33 ($SD = .32$). Power for

576 the random effects meta-analytic effect size with outliers was .05 ($SD = .00$) and without

577 outliers was .05 ($SD = .00$). Unfortunately, power was around 5% for both random effects

578 models with and without outliers. In these studies, 18.9% achieved adequate power of 80%

579 on their found effect size, while 0.0% achieved 80% power for our random effects model with

580 outliers. Finally, without outliers, 0.0% achieved 80% power.

581 **Other Meta-Analytic Estimates.**   We exert caution in interpreting *p*-curve and

582 *p*-uniform analyses on QOL outcomes with and without outliers due to heterogeneity. As

583 seen in Table 1, *p*-uniform estimates were stronger and positive than other techniques

584 because of the high degree of heterogeneity recently described. *p*-curve pictures can be found

585 at the following OSF Link: https://osf.io/4mjqt. Eight studies were significant at $\alpha < .05$,

586 and the studies indicated evidentiary value, $p = .004$. *p*-uniform analyses did not indicate

587 publication bias, $Z = -2.75$, $p = .997$. In PET-PEESE analyses, we found that the intercept

588 was not significant, and therefore, PET was a better estimator of the meta-analytic effect.

589 Table 1 indicates that both of these analyses estimate the effect size around zero, with a

590 confidence interval that includes zero. Selection models correspondingly show small effects

crossing zero, except for random effects models with outliers, that appear to be heavily influenced by the outliers. Trim and fill models are shown in Table 3, and figures are included online. No studies were imputed for these analyses, and therefore, the effect size estimates match the original meta-analysis. Overall, these results appear to point to no effects, ranging across zero with several negative estimates. Interestingly, the correlation of effect sizes across measurement times with outliers was $r = -.37$, 95% CI $[-.62, -.05]$, $t(35) = -2.33$, $p = .026$ and $r = -.64$, 95% CI $[-.80, -.39]$, $t(33) = -4.75$, $p < .001$ without outliers. The effect of expressive writing appears to be positive at short time intervals and decreases into negative effects at longer time intervals.

Discussion

In examining pre- to post-test comparisons across each variable separately, we found that PTS studies indicated a small effect size across all meta-analytic estimates. Both QOL and PTG studies indicated a negligible to small effect size using random effects models. Although the PTG effect in our overall meta-analysis estimate was significant, other methods indicate this small effect is likely not different from zero. QOL was not different from zero, which suggests no effect of expressive writing on quality of life. Interestingly, these results are in contrast to Sloan et al. (2011), which suggested that only certain groups may respond to these tasks. Potentially, the high heterogeneity may be due to the mixed levels of PTSD in these studies, as Di Blasio et al. (2015) indicates that only certain levels of PTSD are responsive to an expressive writing condition.

Expressive writing does not appear to play an important role in influencing positive growth or improved quality of life post task. Ineffective emotional expression may be a contributing factor. In line with this observation, the authors note several limitations. If participants/clients are not deeply engaged with the material, an expressive writing task may not be effective, as Pennebaker and Graybeal (2001) imply that connectedness is an important factor for the task. However, it may be difficult to implement a check for

engagement in these types of research designs. Doing so may also set a context that will inhibit emotional processing and general responses. Research on expressive writing has found a wide range of outcomes for different variables (Frattaroli, 2006), and these various results may explain the large heterogeneity found in this study. Encouragingly, we did not find much evidence of publication bias, and therefore, these estimates may represent a true population effect size. Regardless, methodology of expressive writing studies is variable, as it is applied in different forms across different contexts. Ideally, it would be possible to control for these varied instructions and protocols. However, this is simply not feasible, as most studies do not use measures that examine how engaged an individual is with the material. As such, this current meta-analysis sought to provide readers with a global effect of expressive writing on the aforementioned outcome variables. More studies are needed to examine potential moderating effects of participant engagement.

We also examined the relationship of time between measurements of the dependent variables and the corresponding effect size to determine if effects change over time. For both PTS and PTG, there was no relationship between effect size and time; yet, PTS indicated a small negative correlation. This correlation was not, however, significant. For QOL studies, a medium to large negative correlation was found. A negative relationship between time and effect size implies that writing tasks were more effective in the initial time points, and effects decreased over longer time spans.

The psychological scientific community has shifted focus to reproducibility and research design in the last several years (Nelson, Simmons, & Simonsohn, 2018), and much of this discussion has focused on adequately powering studies for publication (Bakker et al., 2016; S. E. Maxwell, Lau, & Howard, 2015). S. E. Maxwell et al. (2015) and Open Science Collaboration (2015) have shown that the "replication crisis" may be attributed to low power in published studies. The power found in the current meta-analysis was very poor, with very few studies reaching the suggested 80% criterion to adequately power their study. This result was the same when considering individual study characteristics or the estimate true

population effect size. Research by Bakker et al. (2016) indicates that researchers' intuitions about power are particularly poor, and many studies could benefit from more informed power analyses. Anderson, Kelley, and Maxwell (2017) recently published a primer on power, with an online application to help with sample size planning for many types of research designs. Additionally, we encourage researchers to report power analyses of studies in order to better understand methodology for replication and reproducibility.

Meta-analyses, while useful tools to pool for population effect sizes, contain various limitations to their usefulness (Van Elk et al., 2015). As mentioned previously, these analyses can be affected by high heterogeneity, which was found in this study (Aert et al., 2016). Selection models have been criticized when using a smaller number of studies (Van Assen et al., 2015), and trim and fill analyses may not always estimate accurate confidence intervals and funnel plots may be biased with heterogeneity (Terrin, Schmid, Lau, & Olkin, 2003). When focusing on improving the psychological sciences, Van Elk et al. (2015) suggest that the reliability and size of effects may be best elucidated by conducting large preregistered studies. This suggestion will also improve the outlook for power in published studies, and projects such as Many Labs can aide in subsidizing large samples (R. A. Klein et al., 2014). Even with limitations, meta-analyses allow researchers to examine the state of a research area, and we find potential with expressive writing on reducing PTS symptoms, and an overall need for better sample size and power planning for studies.

References

Aert, R. C. M. van, Wicherts, J. M., & Van Assen, M. A. L. M. (2016). Conducting meta-analyses based on p-values: Reservations and recommendations for applying p-uniform and p-curve. *Perspectives on Psychological Science*, *11*(5), 713–729. doi:10.1017/CBO9781107415324.004

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders.* doi:10.1176/appi.books.9780890425596.744053

Anderson, S. F., Kelley, K., & Maxwell, S. E. (2017). Sample-size planning for more accurate statistical power: A method adjusting sample effect sizes for publication bias and uncertainty. *Psychological Science*, *28*(11), 1547–1562. doi:10.1177/0956797617723724

Aslam, N., & Kamal, A. (2013). Gender differences in distress responses, rumination patterns, perceived social support and posttraumatic growth among flood affected individuals. *Journal of Pakistan Psychiatric Society*, *10*, 86–90.

Bakker, M., Hartgerink, C. H. J., Wicherts, J. M., & Maas, H. L. J. van der. (2016). Researchers' intuitions about power in psychological research. *Psychological Science*, *27*(8), 1069–1077. doi:10.1177/0956797616647519

Bodor, N. Z. (2002). *The health effects of emotional disclosure for individuals with Type 1 diabetes* (PhD thesis No. 10-B). Retrieved from http://ezproxy.lib.utexas.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true{\&}db

Borenstein, M., Hedges, L. V., & Rothstein, H. (2007). Meta-analysis fixed effect vs. random effects. Retrieved from https://www.meta-analysis.com/downloads/Meta-analysis fixed effect vs random effects 072607.pdf

Brounéus, K. (2010). The trauma of truth telling: Effects of witnessing in the Rwandan Gacaca Courts on psychological health. *Journal of Conflict Resolution*, *54*(3), 408–437. doi:10.1177/0022002709360322

Bruns, S. B., & Ioannidis, J. P. A. (2016). P-curve and p-hacking in observational research.

PLoS ONE, *11*(2). doi:10.1371/journal.pone.0149144

Buchanan, E. M., Valentine, K. D., & Scofield, J. E. (2017). MOTE. Retrieved from https://github.com/doomlab/MOTE

Carter, E. C., & McCullough, M. E. (2014). Publication bias and the limited strength model of self-control: Has the evidence for ego depletion been overestimated? *Frontiers in Psychology*, *5*(July), 1–11. doi:10.3389/fpsyg.2014.00823

Champely, S. (2016). pwr: Basic functions for power analysis. R package version 1.2-0. Retrieved from https://cran.r-project.org/package=pwr

Cobb, A. R., Tedeschi, R. G., Calhoun, L. G., & Cann, A. (2006). Correlates of posttraumatic growth in survivors of intimate partner violence. *Journal of Traumatic Stress*, *19*(6), 895–903. doi:10.1002/jts.20171

Coburn, K. M., & Vevea, J. L. (2017). Weightr. Retrieved from https://cran.r-project.org/web/packages/weightr/index.html

Cochran, W. G. (1954). Some methods for strengthening the common χ 2 tests. *Biometrics*, *10*(4), 417–451. doi:10.2307/3001616

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Earlbaum.

Cooper, H., Hedges, L. V., & Valentine, J. (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). New York, NY: Russell Sage Foundation.

Costanza, R., Fisher, B., Ali, S., Beer, C., Bond, L., Boumans, R., . . . Snapp, R. (2007). Quality of life: An approach integrating opportunities, human needs, and subjective well-being. *Ecological Economics*, *61*(2-3), 267–276. doi:10.1016/j.ecolecon.2006.02.023

Crespo, M., & Gomez, M. M. (2016). Diagnostic concordance of DSM-IV and DSM-5 posttraumatic stress disorder (PTSD) in a clinical sample. *Psicothema*, *28*(2), 161–166. doi:10.7334/psicothema2015.213

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals,*

716   *and meta-analysis.* New York, NY: Routledge.

717   DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical*

718   *Trials*, *7*(3), 177–188. doi:10.1016/0197-2456(86)90046-2

719   Di Blasio, P., Camisasca, E., Caravita, S. C. S., Ionio, C., Milani, L., Valtolina, G. G., …

720   Valtolina, G. G. (2015). The effects of expressive writing on postpartum depression

721   and posttraumatic stress symptoms. *Psychological Reports*, *117*(3), 856–882.

722   doi:10.2466/02.13.PR0.117c29z3

723   Dursun, P., Steger, M. F., Bentele, C., & Schulenberg, S. E. (2016). Meaning and

724   posttraumatic growth among survivors of the September 2013 Colorado floods.

725   *Journal of Clinical Psychology*, *72*(12), 1247–1263. doi:10.1002/jclp.22344

726   Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing

727   and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455–463.

728   doi:10.1111/j.0006-341X.2000.00455.x

729   Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis

730   detected by a simple, graphical test. *British Medical Journal*, *315*(7109), 629–634.

731   doi:10.1136/bmj.316.7129.469

732   Esterling, B. A., Antoni, M. H., Kumar, M., & Schneiderman, N. (1990). Emotional

733   repression, stress disclosure responses, and Epstein-Barr viral capsid antigen titers.

734   *Psychosomatic Medicine*, *52*, 397–410. doi:10.1097/00006842-199007000-00002

735   Fawzy, N. W., Fawzy, N. W., Hyun, C. S., Elashoff, R., Guthrie, D., Fahey, J. L., & Morton,

736   D. L. (1993). Malignant melanoma. Effects of an early structured psychiatric

737   intervention, coping, and affective state on recurrence and survival 6 years later.

738   *Archives of General Psychiatry*, *50*(9), 681–689.

739   doi:10.1001/archpsyc.1993.01820210015002

740   Francis, G. (2012). Publication bias and the failure of replication in experimental psychology.

741   *Psychonomic Bulletin & Review*, *19*(6), 975–991. doi:10.3758/s13423-012-0322-y

742   Francis, G. (2014). The frequency of excess success for articles in Psychological Science.

Psychonomic Bulletin & Review, 21(5), 1180–1187. doi:10.3758/s13423-014-0601-x

Francis, M. E., & Pennebaker, J. W. (1992). Putting stress into words: The impact of writing on physiological, absentee, and self-reported emotional well-being measures. *American Journal of Health Promotion, 6*(4), 280–287. doi:10.4278/0890-1171-6.4.280

Frankl, V. (1959). *Man's search for meaning* (3rd ed.). Boston, MA: Beacon Press.

Frattaroli, J. (2006). Experimental disclosure and its moderators: A meta-analysis. *Psychological Bulletin, 132*(6), 823–865. doi:10.1037/0033-2909.132.6.823

Frisina, P. G., Borod, J. C., & Lepore, S. J. (2004). A meta-analysis of the effects of written emotional disclosure on the health outcomes of clinical populations. *The Journal of Nervous and Mental Disease, 192*(9), 629–634. doi:10.1097/01.nmd.0000138317.30764.63

Gentes, E. L., Dennis, P. A., Kimbrel, N. A., Rissling, M. B., Beckham, J. C., & Calhoun, P. S. (2014). DSM-5 posttraumatic stress disorder: Factor structure and rates of diagnosis. *Journal of Psychiatric Research, 59*(1), 60–67. doi:10.1016/j.jpsychires.2014.08.014

Gidron, Y., Peri, T., Connolly, J. F., & Shalev, A. Y. (1996). Written disclosure in posttraumatic stress disorder: Is it beneficial for the patient? *The Journal of Nervous and Mental Disease, 184*(8), 505–506. doi:10.1097/00005053-199608000-00009

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5*(10), 3–8. doi:10.3102/0013189X005010003

Goldstein, H. S., Edelberg, R., Meier, C. F., & Davis, L. (1988). Relationship of resting blood pressure and heart rate to experienced anger and expressed anger. *Psychosomatic Medicine, 50*(4), 321–329. doi:10.1097/00006842-198807000-00001

Greenberg, M. A., & Stone, A. A. (1992). Emotional disclosure about traumas and its relation to health: Effects of previous disclosure and trauma severity. *Journal of Personality and Social Psychology, 63*, 75–84. doi:10.1037/h0090372

Gross, J. J., & Levenson, R. W. (1997). Hiding feelings: The acute effects of inhibiting

[770] negative and positive emotion. *Journal of Abnormal Psychology, 106*(1), 95–103.

[771]   doi:10.1037/0021-843X.106.1.95

[772] Harris, A. H. S., Thoresen, C. E., Humphreys, K., & Faul, J. (2005). Does writing affect

[773]   asthma? A randomized trial. *Psychosomatic Medicine, 67*(1), 130–136.

[774]   doi:10.1097/01.psy.0000146345.73510.d5

[775] Hedges, L. V. (1982). Estimation of effect size from a series of independent experiments.

[776]   *Psychological Bulletin, 92*(2), 490–499. doi:10.1037/0033-2909.92.2.490

[777] Henry, E. A., Schlegel, R. J., Talley, A. E., Molix, L. A., & Bettencourt, B. A. (2010). The

[778]   feasibility and effectiveness of expressive writing for rural and urban breast cancer

[779]   survivors. *Oncology Nursing Forum, 37*(6), 749–757. doi:10.1188/10.ONF.749-757

[780] Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring

[781]   inconsistency in meta-analyses. *British Medical Journal, 327*(7414), 557–560.

[782]   doi:10.1136/bmj.327.7414.557

[783] Hilgard, J. (2016). PETPEESE. GitHub. Retrieved from

[784]   https://github.com/Joe-Hilgard/PETPEESE

[785] Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing

[786]   heterogeneity in meta-analysis: Q statistic or I^2 index? *Psychological Methods, 11*(2),

[787]   193–206. doi:10.1037/1082-989X.11.2.193

[788] John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable

[789]   research practices with incentives for truth telling. *Psychological Science, 23*(5),

[790]   524–532. doi:10.1177/0956797611430953

[791] Kelley, K. (2007). Confidence intervals for standardized effect sizes. *Journal of Statistical

[792]   Software, 20*(8), 1–24. doi:10.18637/jss.v020.i08

[793] Klein, R. A., Ratliff, K. A., Vianello, M., Adams, R. B., Bahník, Š., Bernstein, M. J., . . .

[794]   Nosek, B. A. (2014). Investigating variation in replicability. *Social Psychology, 45*(3),

[795]   142–152. doi:10.1027/1864-9335/a000178

[796] Klump, M. C. (2008). Posttraumatic stress disorder and sexual assault in women. *Journal of*

College Student Development, *8225* (May 2014), 37–41. doi:10.1300/J035v21n02

Kross, E., & Ayduk, O. (2011). Making meaning out of negative experiences by self-distancing. *Current Directions in Psychological Science*, *20*(3), 187–191. doi:10.1177/0963721411408883

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*. doi:10.3389/fpsyg.2013.00863

Lancaster, S. L., Klein, K. P., & Heifner, A. (2015). The validity of self-reported growth after expressive writing. *Traumatology*, *21*(4), 293–298. doi:10.1037/trm0000052

Larson, D. G., & Chastain, R. L. (1990). Self-concealment: Conceptualization, measurement, and health implications. *Journal of Social and Clinical Psychology*, *9*(4), 439–455. doi:10.1521/jscp.1990.9.4.439

Lieberman, M. A., & Goldstein, B. A. (2006). Not all negative emotions are equal: The role of emotional expression in online support groups for women with breast cancer. *Psycho-Oncology*, *15*(2), 160–168. doi:10.1002/pon.932

Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does "failure to replicate" really mean? *American Psychologist*, *70*(6), 487–498. doi:10.1037/a0039400

McShane, B. B., Böckenholt, U., & Hansen, K. T. (2016). Adjusting for publication bias in meta-analysis. *Perspectives on Psychological Science*, *11*(5), 730–749. doi:10.1177/1745691616662243

Meshberg-Cohen, S., Svikis, D., & McMahon, T. J. (2014). Expressive writing as a therapeutic process for drug-dependent women. *Substance Abuse*, *35*(1), 80–88. doi:10.1080/08897077.2013.805181

Mogk, C., Otte, S., Reinhold-Hurley, B., & Kröner-Herwig, B. (2006). Health effects of expressive writing on stressful or traumatic experiences - a meta-analysis.

823 *Psychosocial Medicine*, *3*, Doc06.

824 Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with

825 repeated measures and independent-groups designs. *Psychological Methods*, *7*(1),

826 105–125. doi:10.1037/1082-989X.7.1.105

827 Nelson, L. D., Simmons, J., & Simonsohn, U. (2018). Psychology's renaissance. *Annual*

828 *Review of Psychology*, *69*(1), 511–534. doi:10.1146/annurev-psych-122216-011836

829 Open Science Collaboration. (2015). Estimating the reproducibility of psychological science.

830 *Science*, *349*(6251), aac4716–aac4716. doi:10.1126/science.aac4716

831 Pennebaker, J. W. (1989). Confession, inhibition, and disease. In L. Berkowitz (Ed.),

832 *Advances in experimental social psychology* (Vol. 22, pp. 211–244). Academic Press.

833 doi:10.1016/S0065-2601(08)60309-3

834 Pennebaker, J. W. (1993). Putting stress into words: Health, linguistic, and therapeutic

835 implications. *Behaviour Research and Therapy*, *31*(6), 539–548.

836 doi:10.1016/0005-7967(93)90105-4

837 Pennebaker, J. W., & Beall, S. K. (1986). Confronting a traumatic event: Toward an

838 understanding of inhibition and disease. *Journal of Abnormal Psychology*, *95*(3),

839 274–281. doi:10.1037//0021-843X.95.3.274

840 Pennebaker, J. W., & Francis, M. E. (1996). Cognitive, emotional, and language processes in

841 disclosure. *Cognition & Emotion*, *10*(6), 601–626. doi:10.1080/026999396380079

842 Pennebaker, J. W., & Graybeal, A. (2001). Patterns of natural language use: Disclosure,

843 personality, and social integration. *Current Directions in Psychological Science*, *10*(3),

844 90–93. doi:10.1111/1467-8721.00123

845 Pennebaker, J. W., Colder, M., & Sharp, L. K. (1990). Accelerating the coping process.

846 *Journal of Personality and Social Psychology*, *58*(3), 528–537.

847 doi:10.1037//0022-3514.58.3.528

848 Pennebaker, J. W., Kiecolt-Glaser, J. K., & Glaser, R. (1988). Disclosure of traumas and

849 immune function: Health implications for psychotherapy. *Journal of Consulting and*

Clinical Psychology, 56(2), 239–245. doi:10.1037/0022-006X.56.2.239

Possemato, K., Ouimette, P., & Geller, P. (2010). Internet-based expressive writing for
kidney transplant recipients: Effects on posttraumatic stress and quality of life.
Traumatology, 16(1), 49–54. doi:10.1177/1534765609347545

Rachman, S. (1980). Emotional processing. *Behaviour Research and Therapy, 18*(1), 51–60.
doi:10.1016/0005-7967(80)90069-8

Reinhold, M., Bürkner, P. C., & Holling, H. (2018). Effects of expressive writing on
depressive symptoms—A meta-analysis. *Clinical Psychology: Science and Practice,
25*(1). doi:10.1111/cpsp.12224

Sánchez-Meca, J., & Marín-Martínez, F. (2008). Confidence intervals for the overall effect
size in random-effects meta-analysis. *Psychological Methods, 13*(1), 31–48.
doi:10.1037/1082-989X.13.1.31

Scheff, T. J. (1979). *Catharsis in healing, ritual, and drama.* Los Angeles: University of
California Press.

Schoutrop, M. J. A., Lange, A., Hanewald, G., Davidovich, U., & Salomon, H. H. (2002).
Structured writing and processing major stressful events: A controlled trial.
Psychotherapy and Psychosomatics, 71(3), 151–157. doi:10.1159/000056282

Schulenberg, S. E., Hutzell, R. R., Nassif, C., & Rogina, J. M. (2008). Logotherapy for
clinical practice. *Psychotherapy, 45*(4), 447–463. doi:10.1037/a0014331

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). p-curve: A key to the file-drawer.
Journal of Experimental Psychology: General, 143(2), 534–547. doi:10.1037/a0033242

Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2015). Better p-curves: Making p-curve
analysis more robust to errors, fraud, and ambitious p-hacking, a reply to Ulrich and
Miller (2015). *Journal of Experimental Psychology: General, 144*(6), 1146–1152.
doi:10.1037/xge0000104

Slavin-Spenny, O. M., Cohen, J. L., Oberleitner, L. M., & Lumley, M. A. (2011). The effects
of different methods of emotional disclosure: Differentiating posttraumatic growth

from stress symptoms. *Journal of Clinical Psychology, 67*(10), 993–1007. doi:10.1002/jclp.20750

Sloan, D. M., Marx, B. P., & Epstein, E. M. (2005). Further examination of the exposure model underlying the efficacy of written emotional disclosure. *Journal of Consulting and Clinical Psychology, 73*(3), 549–554. doi:10.1037/0022-006X.73.3.549

Sloan, D. M., Marx, B. P., & Greenberg, E. M. (2011). A test of written emotional disclosure as an intervention for posttraumatic stress disorder. *Behaviour Research and Therapy, 49*(4), 299–304. doi:10.1016/j.brat.2011.02.001

Sloan, D. M., Marx, B. P., Bovin, M. J., Feinstein, B. A., & Gallagher, M. W. (2012). Written exposure as an intervention for PTSD: A randomized clinical trial with motor vehicle accident survivors. *Behaviour Research and Therapy, 50*(10), 627–635. doi:10.1016/j.brat.2012.07.001

Sloan, D. M., Marx, B. P., Epstein, E. M., & Lexington, J. M. (2007). Does altering the writing instructions influence outcome associated with written disclosure? *Behavior Therapy, 38*(2), 155–168. doi:10.1016/j.beth.2006.06.005

Smithson, M. (2001). Correct confidence intervals for various regression effect sizes and parameters: The importance of noncentral distributions in computing intervals. *Educational and Psychological Measurement, 61*(4), 605–632. doi:10.1177/00131640121971392

Smithson, M. (2003). *Confidence intervals.* Thousand Oaks, CA: Sage.

Smyth, J. M. (1998). Written emotional expression: Effect sizes , outcome types, and moderating variables. *Journal of Consulting and Clinical Psychology, 66*(1), 174–184. doi:10.1037/0022-006X.66.1.174

Smyth, J. M., Stone, A. A., Hurewitz, A., & Kaell, A. (1999). Effects of writing about stressful experiences on symptom reduction in patients with asthma or rheumatoid arthritis: A randomized trial. *JAMA: The Journal of the American Medical*

903 *Association*, *281*(14), 1304–1309. doi:10.1001/jama.281.14.1304

904 Stanley, T. D. (2005). Beyond publication bias. *Journal of Economic Surveys*, *19*(3),

905 309–345. doi:10.1111/j.0950-0804.2005.00250.x

906 Stanley, T. D., & Doucouliagos, H. (2014). Meta-regression approximations to reduce

907 publication selection bias. *Research Synthesis Methods*, *5*(1), 60–78.

908 doi:10.1002/jrsm.1095

909 Stanton, A. L., Danoff-Burg, S., Sworowski, L. A., Collins, C. A., Branstetter, A. D.,

910 Rodriguez-Hanley, A., . . . Austenfeld, J. L. (2002). Randomized, controlled trial of

911 written emotional expression and benefit finding in breast cancer patients. *Journal of*

912 *Clinical Oncology*, *20*(20), 4160–4168. doi:10.1200/JCO.2002.08.521

913 Taku, K., Calhoun, L. G., Cann, A., & Tedeschi, R. G. (2008). The role of rumination in the

914 coexistence of distress and posttraumatic growth among bereaved Japanese

915 University students. *Death Studies*, *32*(5), 428–444. doi:10.1080/07481180801974745

916 Tedeschi, R. G., & Calhoun, L. G. (1995). *Trauma & transformation: Growing in the*

917 *aftermath of suffering.* Thousand Oaks, CA: Sage Publications.

918 Tedeschi, R. G., & Calhoun, L. G. (2004). Posttraumatic growth: Conceptual foundations

919 and empirical evidence. *Psychological Inquiry*, *15*(1), 1–18.

920 doi:10.1207/s15327965pli1501

921 Terrin, N., Schmid, C. H., Lau, J., & Olkin, I. (2003). Adjusting for publication bias in the

922 presence of heterogeneity. *Statistics in Medicine*, *22*(13), 2113–2126.

923 doi:10.1002/sim.1461

924 Theofilou, P. (2013). Quality of life: Definition and measurement. *Europe's Journal of*

925 *Psychology*, *9*(1), 150–162. doi:10.5964/ejop.v9i1.337

926 Van Aert, R. C. M. (2017). P-uniform. GitHub. Retrieved from

927 https://github.com/RobbievanAert/puniform

928 Van Assen, M. A. L. M., Van Aert, R. C. M., & Wicherts, J. M. (2015). Meta-analysis using

929 effect size distributions of only statistically significant studies. *Psychological Methods*,

20(3), 293–309. doi:http://dx.doi.org/10.1037/met0000025

Van Elk, M., Matzke, D., Gronau, Q. F., Guan, M., Vandekerckhove, J., & Wagenmakers, E.-J. (2015). Meta-analyses are no substitute for registered replications: A skeptical perspective on religious priming. *Frontiers in Psychology*, *6*, 1365. doi:10.3389/fpsyg.2015.01365

Van Emmerik, A. A. P., Reijntjes, A., & Kamphuis, J. H. (2013). Writing therapy for posttraumatic stress: A meta-analysis. *Psychotherapy and Psychosomatics*, *82*(2), 82–88. doi:10.1159/000343131

Vevea, J. L., & Hedges, L. V. (1995). A general linear model for estimating effect size in the presence of publication bias. *Psychometrika*, *60*(3), 419–435. doi:10.1007/BF02294384

Vevea, J. L., & Woods, C. M. (2005). Publication bias in research synthesis: Sensitivity analysis using a priori weight functions. *Psychological Methods*, *10*(4), 428–443. doi:10.1037/1082-989X.10.4.428

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. doi:10.18637/jss.v036.i03

Walker, B. L., Nail, L. M., & Croyle, R. T. (1999). Does emotional expression make a difference in reactions to breast cancer? *Oncology Nursing Forum*, *26*(6), 1025–1032.

Wang, X., Gao, L., Shinfuku, N., Zhang, H., Zhao, C., & Shen, Y. (2000). Longitudinal study of earthquake-related PTSD in a randomly selected community sample in North China. *American Journal of Psychiatry*, *157*(8), 1260–1266. doi:10.1176/appi.ajp.157.8.1260

Weiss, T. (2002). Posttraumatic growth in women with breast cancer and their husbands – An intersubjective validation study. *Journal of Psychosocial Oncology*, *20*(2), 65–80. doi:10.1300/J077v20n02_04

Wilson, K. G., & DuFrene, T. (2009). *Mindfulness for two: An acceptance and commitment Therapy approach to mindfulness in psychotherapy.* Oakland, CA: New Harbinger

956   Publications.

957 Yilmaz, M., & Zara, A. (2016). Traumatic loss and posttraumatic growth: The effect of

958   traumatic loss related factors on posttraumatic growth. *Anatolian Journal of*

959   *Psychiatry, 17*(1), 5–11. doi:10.5455/apd.188311

Table 1

Effect Size Estimates for PTS Results

Model	Fixed Effects	Random Effects	Fixed No Outliers	Random No Outliers
Overall Effects	0.34 [0.31, 0.37]	0.39 [0.32, 0.46]	0.32 [0.29, 0.35]	0.36 [0.29, 0.42]
Z Values	21.75, $p < .001$	11.06, $p < .001$	20.00, $p < .001$	11.03, $p < .001$
p-Uniform	0.60 [0.50, 0.71]	-	0.57 [0.47, 0.67]	-
PET	0.12 [0.03, 0.21]	-	0.11 [0.02, 0.20]	-
PEESE	0.25 [0.20, 0.30]	-	0.23 [0.18, 0.28]	-
Selection Models	0.33 [0.28, 0.37]	0.45 [0.33, 0.57]	0.29 [0.24, 0.33]	0.39 [0.27, 0.50]
Trim and Fill	0.26 [0.23, 0.29]	0.26 [0.18, 0.34]	0.25 [0.22, 0.28]	0.25 [0.18, 0.32]

Note. [] indicates the 95 percent confidence interval for each effect size estimate.

Table 2

Effect Size Estimates for PTG Results

Model	Fixed Effects	Random Effects
Overall Effects	0.10 [0.02, 0.17]	0.10 [0.02, 0.17]
Z Values	2.45, *p* = .014	2.45, *p* = .014
p-Uniform	-0.11 [-1.43, 0.42]	-
PET	0.06 [-0.20, 0.32]	-
PEESE	0.08 [-0.04, 0.20]	-
Selection Models	0.09 [-0.01, 0.18]	0.09 [-0.03, 0.20]
Trim and Fill	0.10 [0.02, 0.17]	0.10 [0.02, 0.17]

Note. [] indicates the 95 percent confidence interval for each effect size estimate.

Table 3

Effect Size Estimates for QOL Results

Model	Fixed Effects	Random Effects	Fixed No Outliers	Random No Outliers
Overall Effects	-0.01 [-0.07, 0.05]	-0.01 [-0.16, 0.13]	-0.01 [-0.07, 0.05]	-0.01 [-0.11, 0.09]
Z Values	-0.33, $p = .745$	-0.18, $p = .860$	-0.25, $p = .805$	-0.20, $p = .838$
p-Uniform	0.79 [0.33, 1.61]	-	0.62 [0.10, 0.96]	-
PET	0.05 [-0.26, 0.36]	-	0.05 [-0.29, 0.38]	-
PEESE	0.00 [-0.17, 0.17]	-	0.00 [-0.19, 0.19]	-
Selection Models	-0.06 [-0.12, 0.01]	0.51 [-0.09, 1.12]	-0.04 [-0.11, 0.03]	0.05 [-0.15, 0.24]
Trim and Fill	-0.01 [-0.07, 0.05]	-0.01 [-0.16, 0.13]	-0.01 [-0.07, 0.05]	-0.01 [-0.11, 0.09]

Note. [] indicates the 95 percent confidence interval for each effect size estimate.
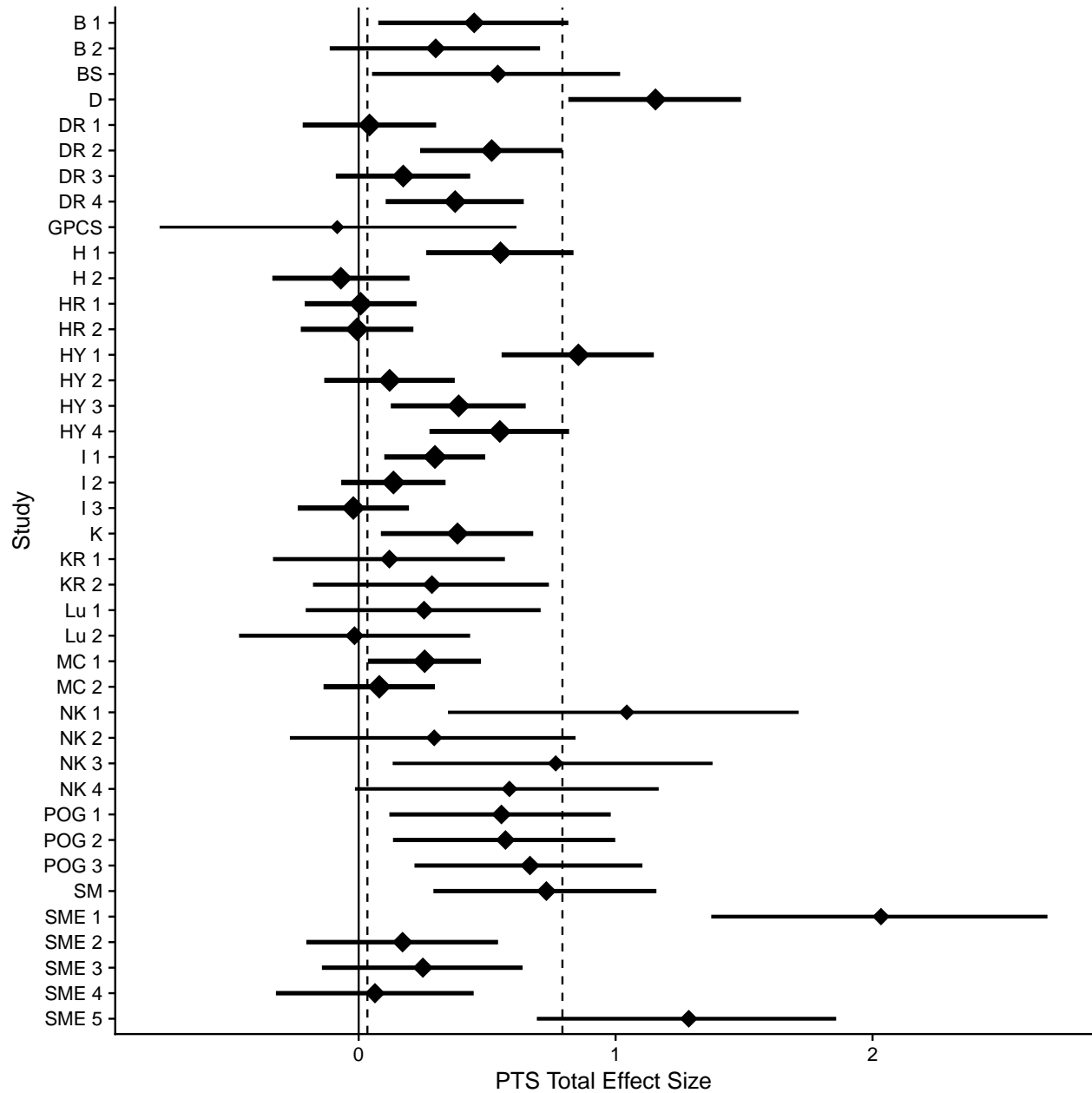
Figure 1. Effect sizes and their non-centralized confidence interval for PTS total scores.

Dashed lines indicated average non-weighted lower and upper confidence interval limits.

Diamond size indicates normalized study weight from a random effects model.
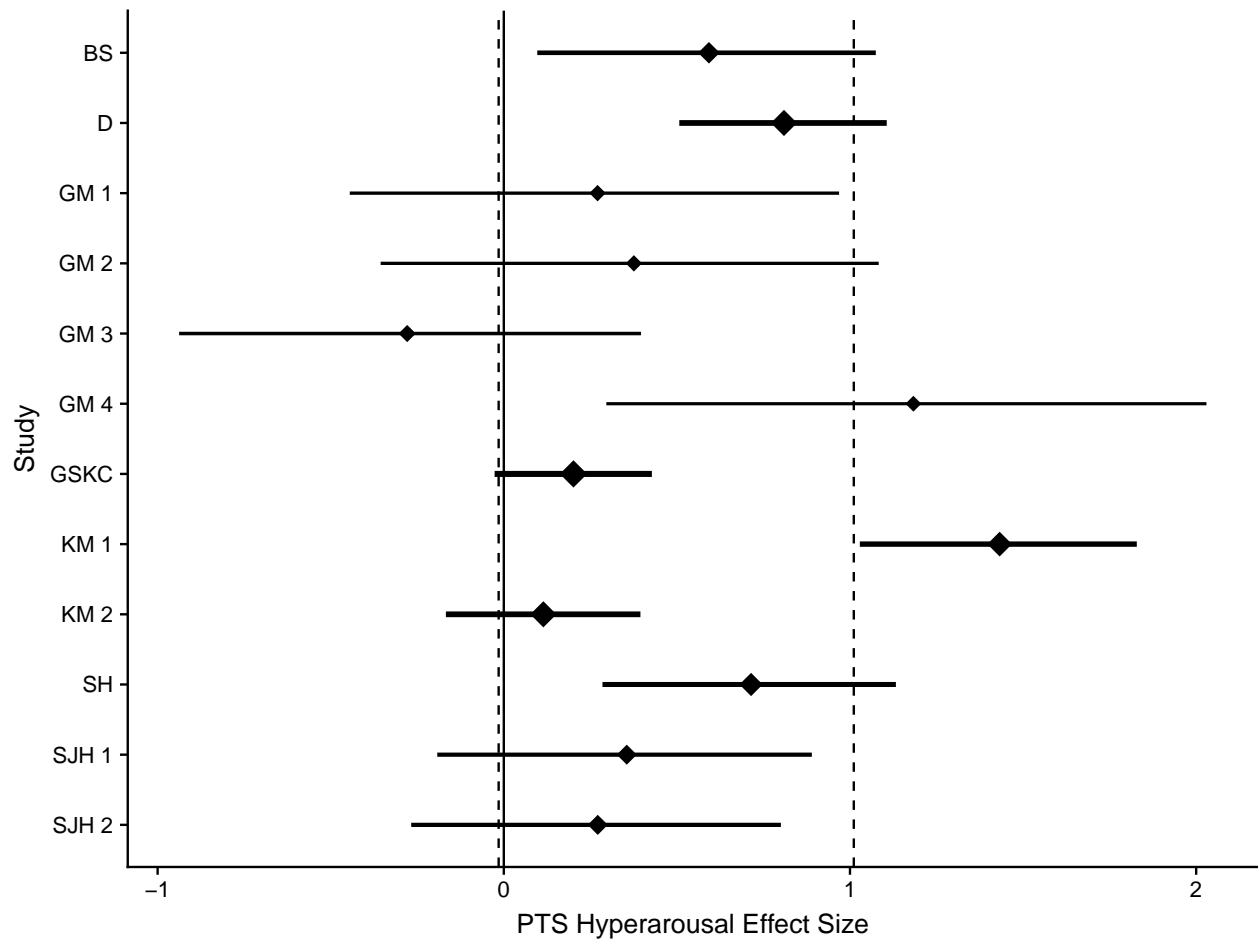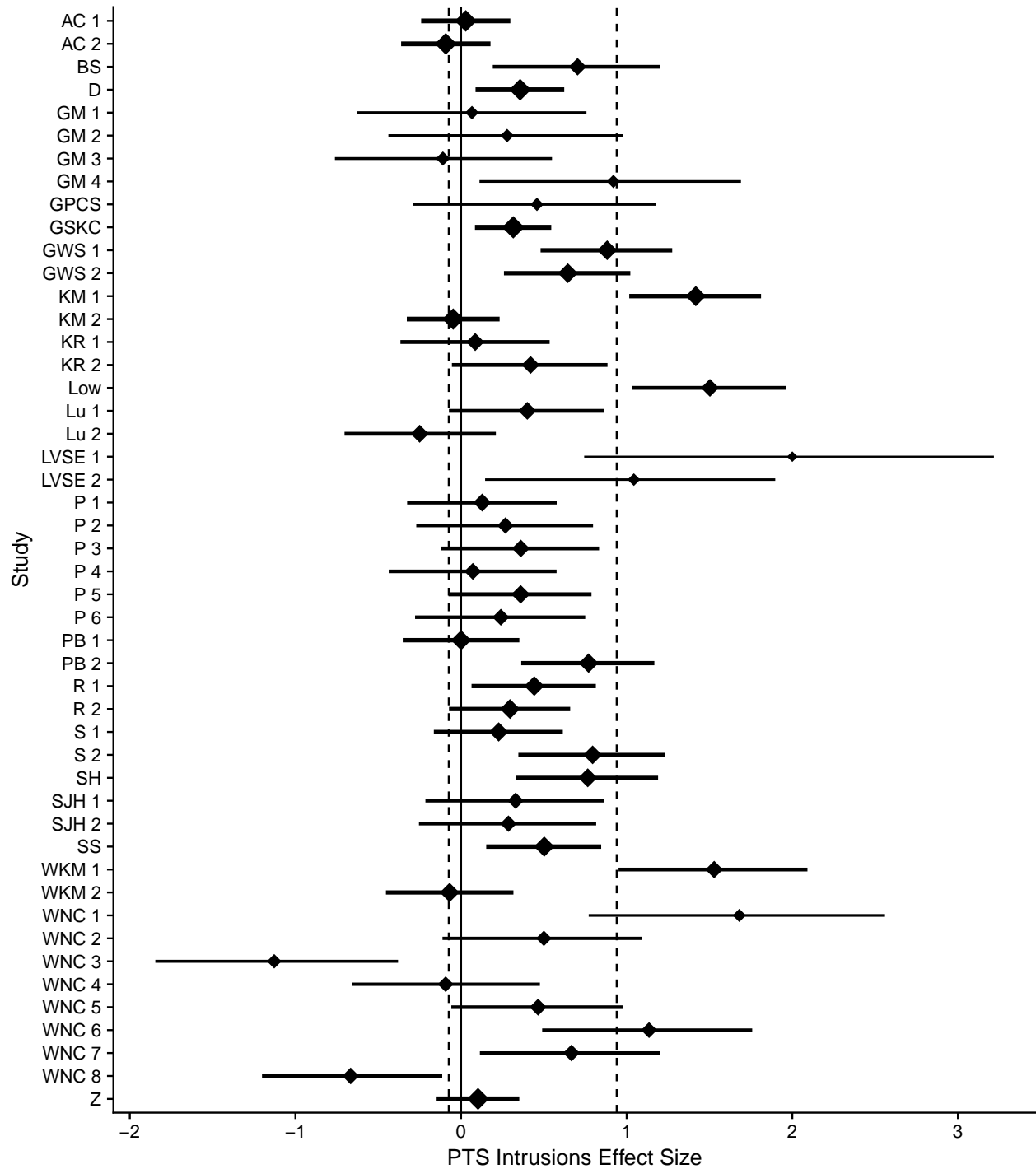
Figure 2. Effect sizes and their non-centralized confidence interval for PTS Hyperarousal. Dashed lines indicated average non-weighted lower and upper confidence interval limits. Diamond size indicates normalized study weight from a random effects model.
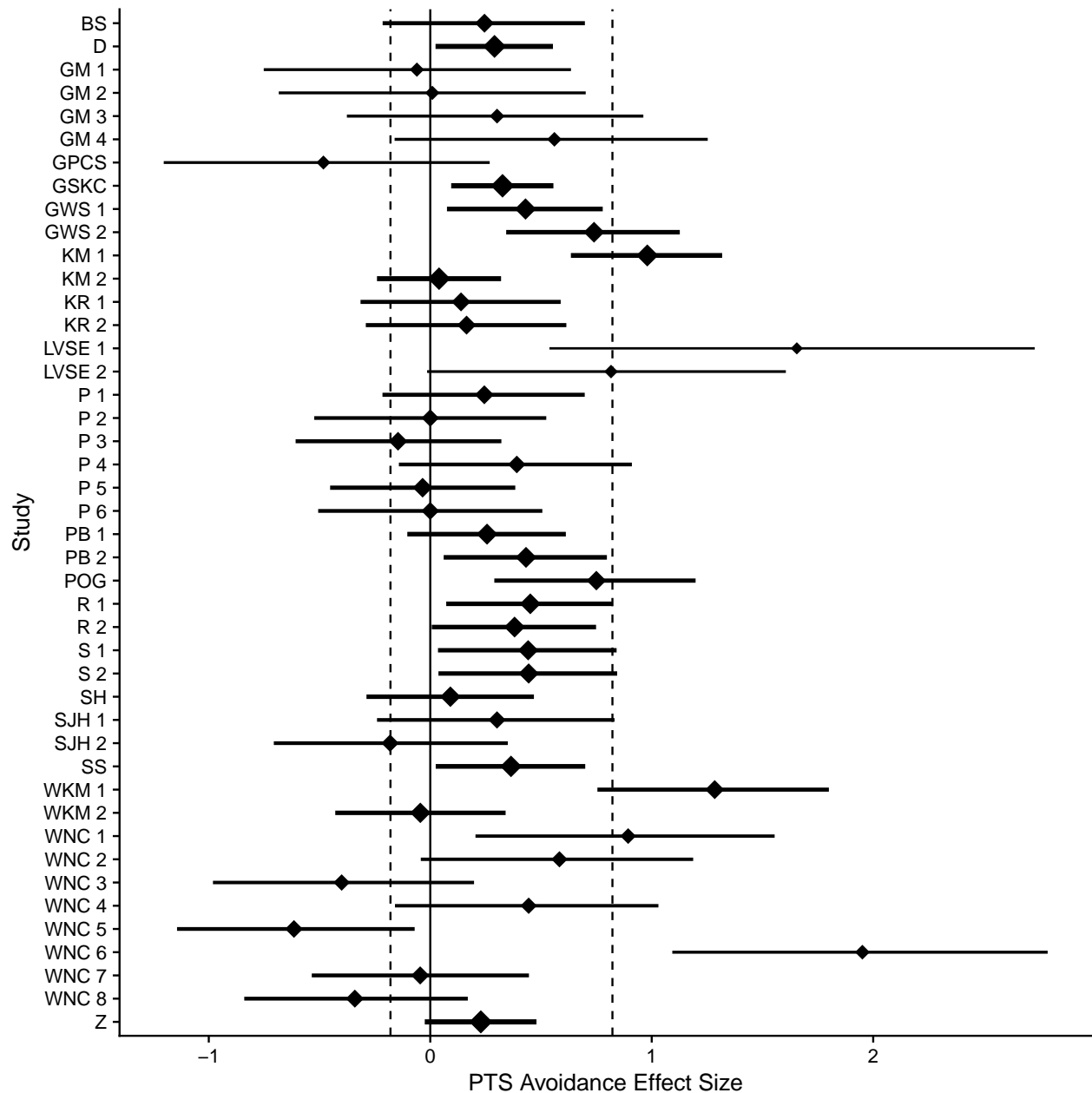
Figure 3. Effect sizes and their non-centralized confidence interval for PTS Intrusion scores. Dashed lines indicated average non-weighted lower and upper confidence interval limits. Diamond size indicates normalized study weight from a random effects model.

Figure 4. Effect sizes and their non-centralized confidence interval for PTS Avoidance Scores. Dashed lines indicated average non-weighted lower and upper confidence interval limits. Diamond size indicates normalized study weight from a random effects model.
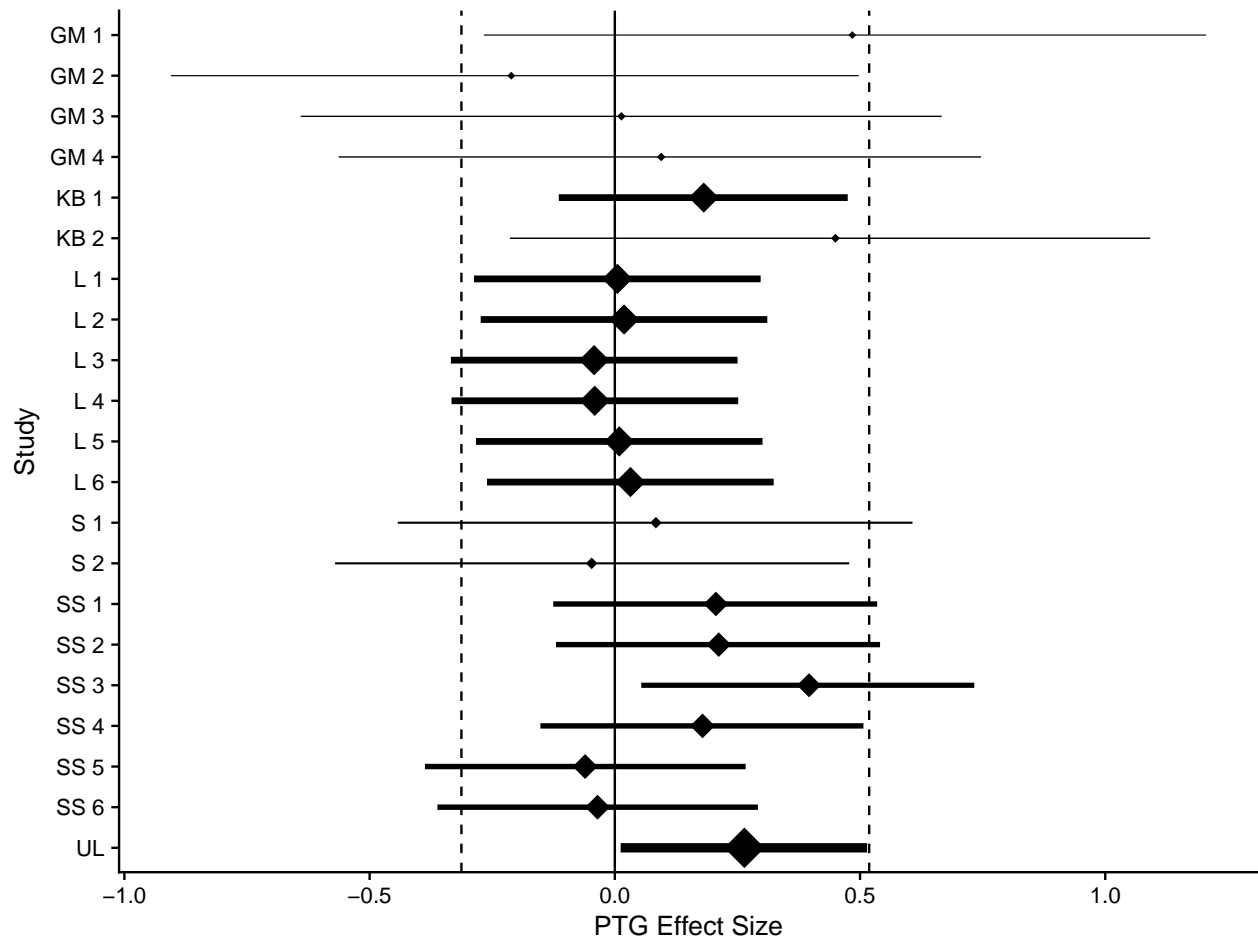
Figure 5. Effect sizes and their non-centralized confidence interval for PTG outcome variables. Dashed lines indicated average non-weighted lower and upper confidence interval limits. Diamond size indicates normalized study weight from a random effects model.
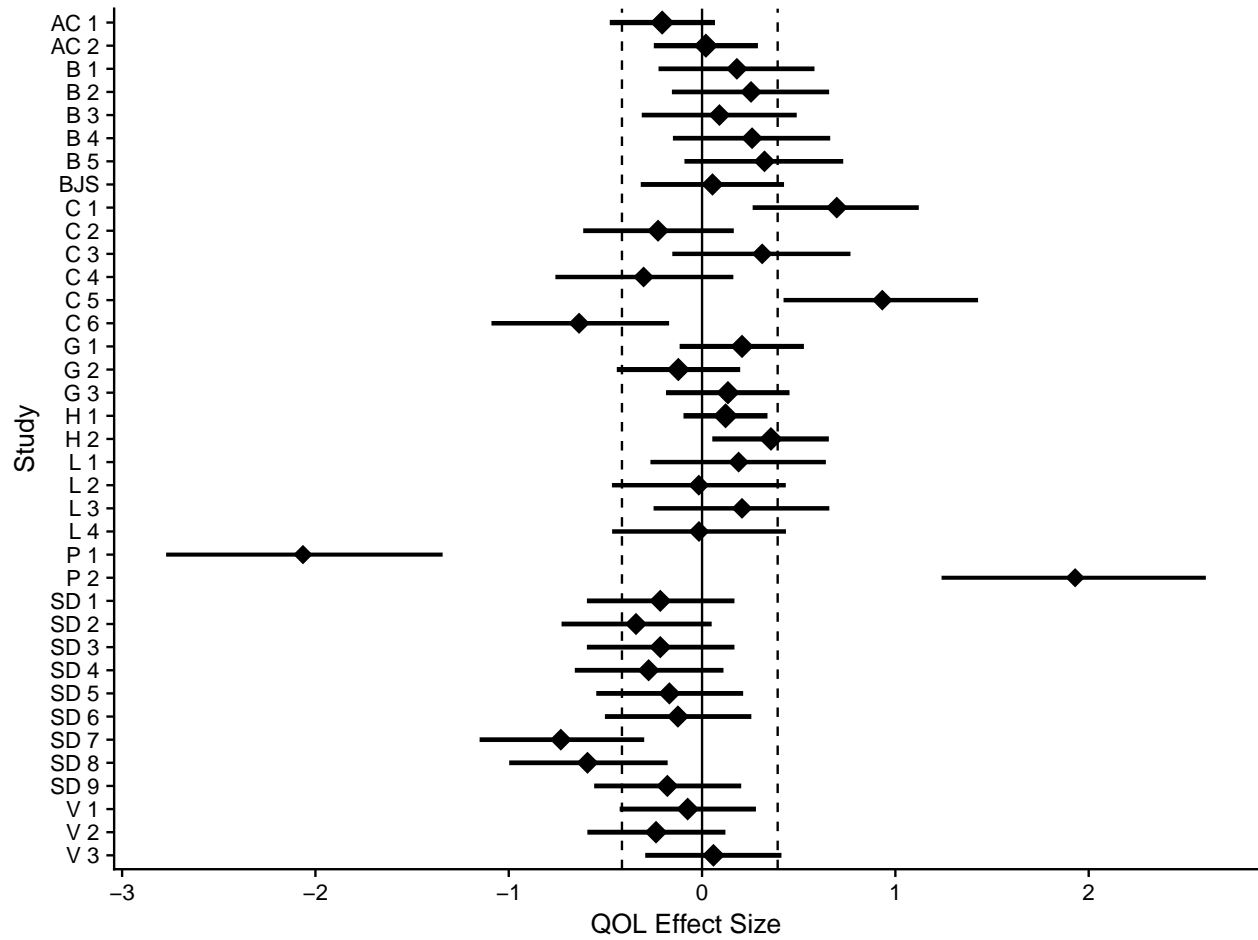
Figure 6. Effect sizes and their non-centralized confidence interval for QOL outcome variables.
Dashed lines indicated average non-weighted lower and upper confidence interval limits.
Diamond size indicates normalized study weight from a random effects model.