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

Jeffrey M. Pavlacic

Erin M. Buchanan

Nicholas P. Maxwell

Missouri State University

Abstract

Repressing negative thoughts or emotions can be detrimental to both physical and psychological health. Additionally, psychological dysfunction can create problematic health behaviors. Clinicians have considered a number ways for individuals suffering from psychological distress to effectively expressive these emotions. One therapeutic approach implemented by clinicians suggests that writing about these negative emotions can lead to reductions in psychological distress. Pennebaker (1997) suggests that writing about a traumatic event can reduce the psychological stress associated with that event. Expressive writing interventions have been adapted to interventions for a multitude of different variables and populations. Results from a meta-analysis suggest a small effect size for the effect of expressive writing on Quality of Life (QOL) and Posttraumatic Growth (PTG) variables for a variety of health-related illnesses and psychological disorders as a result of expressive writing. Additional research utilizing an expressive writing intervention under Pennebaker’s paradigm is necessary in order to discover the effectiveness of expressive writing on different health-related and psychological diagnoses, as the few studies published in the literature were found to be underpowered to detect significant changes in outcomes.

*Keywords: expressive writing, posttraumatic growth, quality of life.*

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

**Problems Associated with Repressing Negative Emotions**

Inhibitory theory discusses how repressing negative thoughts or emotions can be detrimental to both physical and psychological health. Furthermore, inhibitory theory postulates that individuals experiencing traumatic events are more likely to repress thoughts and feelings about a given traumatic experience. These repressive maneuvers have the capability to lead to social concerns and overall psychological dysfunction (Pennebaker, 1989; Pennebaker & Beall, 1986). Psychological dysfunction can have detrimental effects on an individual’s health, including unhealthy everyday life habits such as low activity levels, lower quality of life, and inability to progress after a traumatic event. These effects on health could lead to biological problems, especially immune system deficiencies and problems with neurotransmitters (Pennebaker & Beall, 1986)

**Expressive Writing**

Clinicians have considered and implemented many different ways for patients to successfully express their emotions. Verbalizing emotions has the capability to improve psychological well-being and improve psychological health, especially after experiencing a traumatic event (Smyth & Pennebaker, 2008). One type of therapeutic approach suggests that writing about negative emotional experiences leads to significant reduction in psychological distress (Pennebaker & Beall, 1986). Pennebaker (1997) explored the use of expressive writing to elucidate the potential benefits of verbal expression of emotion. Pennebaker’s research suggests that writing about a traumatic event can help to decrease psychological distress related to the event. Since Pennebaker started utilizing expressive writing, many variations have been adapted to examine the effectiveness of different types of expressive writing for different symptoms and psychopathology (Manier & Olivares, 2005). Over 200 studies have been conducted utilizing expressive writing in some form. Yet, the effectiveness of expressive writing remains unclear. Baikie and Wilhelm (2005) posit that expressive writing leads to improvements in both physical and psychological health across both clinical and non-clinical populations. Furthermore, writing about a previous traumatic experience leads to a short-term increase in arousal as well as a decrease in chronic health problems (Pennebaker & Beall, 1986). Usually, studies of this nature involve participants writing about a traumatic event for three to five sessions for a time period of 15-20 minutes with a pre- to post-writing intervention comparison as the point of interest. Interestingly, at long-term follow up, studies have found evidence of health benefits from expressive writing, such as greater psychological wellbeing, reduced depressive symptoms prior to examination, and fewer cases of posttraumatic intrusion and avoidance symptoms (Baikie & Wilhelm, 2005). However, a consensus on the effectiveness of expressive writing has not yet been established on various outcome variables.

**Why is a Meta-Analysis Necessary?**

The literature shows contradictory results of the effectiveness of this type of intervention, which may be a factor as to why clinicians are hesitant to incorporate expressive writing into their treatment program. Henry, Schlegel, Talley, Molix, and Bettencourt (2010) found that expressive writing only benefited a rural population for those individuals surviving breast cancer. Lancaster, Klein, and Heifner (2015) found no significant evidence that expressive writing can be considered an effective approach. Expressive writing tasks fit well within the framework of different psychological interventions and can be adapted for treatment, which is why the literature includes many different studies looking at a multitude of variables. However, it is important to focus on individual variables in order to determine the effectiveness of expressive writing for specific diagnoses and psychopathology. As previously mentioned, some studies have found long-term benefits of expressive writing on psychological well-being (Park & Blumberg, 2002). However, other studies, such as the research completed by Lancaster et al. (2015), have found no evidence supporting the utilization of expressive writing as an effective therapeutic approach. Thus, it is necessary to evaluate the effectiveness of expressive writing on specific outcome variables, and we chose to focus specifically on posttraumatic growth and quality of life, in line with the current positive psychology trend.

**Posttraumatic Stress Disorder and Posttraumatic Growth**

Posttraumatic Stress Disorder (PTSD) is a disorder involving reoccurring thoughts or experiences after a traumatic event or experience. The diagnosis is based on 20 symptoms structured into four different subsets. These subsets are as follows: re-experiencing, avoidance, negative alterations in cognition and mood, and arousal (Crespo & Gomez, 2016). Research conducted on the effectiveness of expressive writing on PTSD symptoms has been less successful and shows outcomes that are not as effective as other studies (Sloan, Marx, & Greenberg, 2011). Posttraumatic growth (PTG) is a positive experience after a traumatic event (Yilmaz & Zara, 2016). Expressive writing has been shown to be an effective method for reducing psychological distress among those suffering from trauma (Sloan, Marx, Epstein, & Lexington, 2007). PTSD is concerning, specifically among Iraq and Afghanistan war veterans (Gentes & Cambone, 2013). It is important to examine the effectiveness of an expressive writing intervention and look at how effective this approach is in promoting PTG and overall psychological health. Speculation suggests that those meeting the criteria for moderate PTSD benefit more from expressive writing interventions as opposed to those with greater PTSD symptoms (Di Blasio et al., 2015). If particular studies utilizing an expressive writing paradigm are shown to benefit the patient and create a more positive lifestyle, then it is necessary to determine the overall effectiveness of expressive writing on PTG.

**Quality of Life**

Quality of Life (QOL) is another variable, related to PTSD and psychology health, that is worth examining with expressive writing interventions. QOL is described as a concept comprised of multiple domains, both subjective and objective. Objectively, QOL is a measure of the extent to which an individual’s needs are met. Subjectively, QOL measures an individual’s attitude towards their given situation (Costanza et al., 2006). Pennebaker and Graybeal (2001) suggested that expressive writing allows one to feel more connected with their surroundings. Furthermore, they explain that expressive writing allows people to see things in a different way and better understand themselves. By understanding a traumatic event, one is able 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, 1984). Higher QOL may be considered a type of PTG, which is why it would be interesting to examine the effectiveness of studies utilizing expressive writing to improve QOL and PTG in the same study to compare effectiveness.

**Purpose of Current Meta-Analysis**

The purpose of this meta-analysis is to examine studies utilizing expressive writing on QOL and PTG variables. Due to inconsistent results in current studies published, it is important to elucidate the effectiveness of expressive writing on promoting positive change after a traumatic event and improving overall quality of life. Thus, a meta-analysis will allow a collected look at the use of expressive writing in these situations. This particular meta-analysis examines studies of patients with different types of psychopathology and medical diagnoses on PTG and QOL outcomes. The main focus is to examine PTG and QOL and the effect sizes related to expressive writing interventions utilizing Pennebaker’s paradigm.

**Method**

**Data Collection**

Studies were collected through online databases, such as PsycINFO and Google Scholar using the following search terms: *Posttraumatic Growth Expressive Writing, Expressive Writing Posttraumatic Growth, PTG Expressive Writing, Expressive Writing PTG, Quality of Life Expressive Writing, Expressive Writing Quality of Life, QOL Expressive Writing, Expressive Writing QOL, Expressive Writing.* Within these articles, the change in PTG and 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 because they received the expressive writing intervention. 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.

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

Each study implemented a pre-test to post-test style repeated measures design, usually with paired *t*-tests, ANOVA, or regression analyses. The means, standard deviations, and *N* values were collected from each study, with the exception of the Smith et al. (2015) wherein a regression coefficient and corresponding *t-*value for time was used. Cohen’s *d* values were calculated using the following formula for paired *t* using means and standard deviations:

Equation 1 is described in detail in Cumming (2013) as an alternative to the traditional calculation of *d* for paired samples *t*, wherein the denominator is the standard deviation of the difference scores (*ddiff*, alternatively *dRM*). Equation 1 for *davg* not only allows for calculations from published articles that do not include *SDdiff* (i.e., most articles included), but also has been shown to be less upwardly biased than *ddiff*. Alternative formulas include controlling for *r* between paired levels, as described in Lakens (2013); however, these values were not available in the selected articles, and *davg*is a recommended effect size for paired designs. The regression coefficient was translated to *d* using equation 2, which is a mathematical translation of the *t* formula (Rosenthal, 1991):

We planned to use traditional and newer methods of meta-analysis, following guidelines from Cooper, Hedges, and Valentine (2009), as well as van Aert, Wicherts, and van Assen (2016). Sampling variance of the effect sizes were estimated using the *escalc* function from the *metafor* package (reference cite the thing). The variance formula was originally published in Morris and DeShon (2002) and is shown in equation 3:

In this formula, *n* is the number of paired observations, *d* is the calculated effect size, and *c* is a correction factor, wherein *df* are *n* – 1 (Hedges, 1982):

We used the *metagen* function in the *metafor* package to calculate both fixed and random effects models, which uses standard error of the effect to calculate overall estimates of an effect and their confidence intervals. Thus, we took the square root of the variance estimate for standard error.

Given these calculations, the goal of this analysis is to calculate a combined effect size, along with a confidence interval for study planning and an assessment of the literature. A fixed effects model requires the assumption that there is a true population effect size across all studies. By including two measures of positive psychological outcomes, this assumption may be tenuous, and therefore, a random effects model was also calculated. In random effects models, the true effect is assumed to vary across studies (Borenstein, Hedges, & Rothstein, 2007). For a fixed effects model, the effect sizes are weighted by their inverse variance (Sanchez-Meca & Marin-Martinez, 2008), which is calculated automatically in *metafor* by:

The advantage to this procedure is that analyses are weighted by their precision, that is, that studies with more information (often, larger samples), are given larger weights in the overall estimated effect size (Borenstein et al., 2007). Random effects models are also weighted by inverse variance, with an additional correction for variance between studies, , as described by DerSimonian and Laird (1986):

Confidence intervals were calculated in two ways for this study. Cumming (2012), Kelley (2007), and Smithson (2003) have shown that the distribution of *d* values are non-normal, and thus, CIs should be estimated using the non-centrality parameter and a non-normal distribituion. These values were calculated using *R* scripts that interatively estimate the appropriate non-centrality parameter and convert back to *d* values (i.e., non-centrality parameter divided by the square root of *n,* Smithson, 2003; Buchanan, cite). However, the *metafor* package in *R* uses central distributions to estimate CIs for each study and overall effect sizes. Therefore, we present both sets of values for the interested reader, as meta-analytic procedure has not quite caught up to our understanding of the distributions of effect sizes.

**Results**

**Overall Effect Size**

Experiments were examined for potential outliers using the *metafor* package, which calculates traditional regression influence values, such as Cook’s and hat values. These values indicate change in overall meta-analytic model with and without the experiment, thus, determining their impact on the pooled effect size. One experiment was a common outlier between fixed effects and random effects estimates (bolded in Table 1). Results from the meta-analysis models are presented with and without the outlier (Table 2). As suggested by van Aert et al. (2016), we also looked for evidence of *p*-hacking, where in researchers may manipulate their data to find significant effects. *p*-hacking was defined as finding *p*-values between .025 and .05 with power lower than 60%. No studies met this criterion, as shown in Table 1.

Table 1 includes all study characteristics including sample size, *d* values with their standard errors, and 95% normal distribution CI estimates. Only two effect sizes were found to be significantly different from zero, as indicated by the *z* and *p*-value columns. These values were calculated directly from estimated effect size, rather than statistics provided in the study, which allows us to focus in on only PTG and QOL variables for intervention groups. Only two (11.8%) studies were found to have significant changes in outcomes. Power is discussed below. Figure 1 is a forest plot of each study’s effect size, along with non-central CIs, calculated as described earlier, which show similar results to Table 1. The box’s size on a forest plot represents the weight of each study, and horizontal lines display the CI. Experiments are grouped by outcome variable, and the unweighted average of the effects and CIs are presented at the bottom of the plot. Overall, we found that the overall effect size was negligible to small, *d* = 0.13, 95%CI [0.05, 0.22], with nearly identical results for random effects models after the exclusion of the outlier study. Both PTG and QOL showed similar sizes, with a slightly larger range of expected effects for QOL. Table 2 indicates all estimates and normal CIs.

**Homogeneity**

A pre-requisite for newer meta-analysis techniques, such as *p*-curve/uniform and PET-PEESE, includes the assessment of homogeneity of the effects (van Aert et al., 2016). Using the *metafor* package, we calculated the *Q-*statistic and the *I­2* index (Cochran, 1954; Huedo-Medina, Sanchez-Meca, & Marin-Martinez, 2006). Both *Q* and *I­2* measure heterogeneity or inconsistency between studies, each with their own pros and cons. *Q* indicates if heterogeneity exists, while *I­2* measures the percentage of heterogeneity, along with a CI. Both can be biased with a small number of expeirments (Higgins, Thompson, Deeks, & Altman, 2003; Huedo-Medina, et al., 2006). The inclusion of the outlier indicated heterogeneity, *Q*(16) = 111.94, *p* < .001, *I2* = 85.7%, 95%CI [78.6%; 90.5%]; however, the removal of that outlier eliminated this effect, *Q*(15) = 17.13, *p* = .31, *I2* = 12.5%, 95%CI [0.0%; 49.8%]. The PTG experiments did not show heterogeneity, *Q*(8) = 3.76, *p* = .88 *I2* = 0.0%, 95%CI [0.0%; 25.2%]. The QOL experiments were heavily influenced by the outlier, *Q*(7) = 102.43, *p* < .001, *I2* = 93.2%, 95%CI [88.8%; 95.8%], and its removal decreased, but did not eliminate heterogeneity, *Q*(6) = 13.29, *p* = .04, *I2* = 54.8%, 95%CI [0.0%; 80.7%]. Because of the heterogeneity present even when smaller subgroups were calculated, we did not calculate *p*-curve/uniform and PET-PEESE estimates, as they are biased for heterogenous effects (van Aert, et al., 2016; Moreno, et al., 2009).

**Power**

Power was calculated in two different manners using the *pwr* package in *R* (Champely, 2009). *Post hoc* power was first calculated using the statistics from the study, namely, sample size and effect size. Second, power was calculated using the study sample size and estimated overall effect *d* = 0.13, as explained by Francis (2012, 2014). The first estimate indicates the likelihood of finding the effect from the study, while the second indicates the likelihood of detecting a true population effect size. The average power for individual study calculations was *M* = .26 (*SD* = .29), and the average power for overall effect calculations was *M* = .12 (*SD* = .06). Only two studies had power greater than 80% given their own study characteristics, which is a common standard in psychology (Cohen, 1988). The calculation of power using the overall effect indicated that all studies were underpowered, with none reaching a “coin flip’s” chance at detecting the effect (Cohen, 1990). Another estimate of bias, the Test of Excessive Success, was also not appropriate for our study, as this test requires one publication to have at least four expeirments (Francis, 2012).

**Discussion**

These results seem to contradict traditional views of expressive writing in that most studies find expressive writing to be effective for improving psychological functioning. Few empirical studies seemed to fit the criteria for this particular study in that many studies did not use Posttraumatic Growth or Quality of Life as an outcome variable. Thus, additional studies need to be conducted in order to assess whether or not an expressive writing intervention using Pennebaker’s paradigm is effective for these particular outcome variables. While many studies have been conducted to assess both the effectiveness and efficacy of expressive writing, it is necessary to conduct studies that examine these particular variables. Doing so could possibly help to decide whether or not expressive writing is a useful intervention for improving quality of life in individuals and those individuals who have suffered a traumatic event. These findings, however, suggest that this particular intervention is not effective for these particular variables.

**Limitations**

Publication bias can’t really be tested or corrected for

**References**

Baikie, K. A., & Wilhelm, K. (2005). Emotional and physical health benefits of expressive writing. *Advances in Psychiatric Treatment*, *11*(11), 338–346. https://doi.org/10.1192/apt.11.5.338

Crespo, M., & Gomez, M. M. (2016). Diagnostic concordance of DSM-IV and DSM-5 Posttraumatic Stress Disorder (PTSD) in a clinical sample, *28*(2), 161–166. https://doi.org/10.7334/psicothema2015.213

Gentes, A., & Cambone, M. (2013). Designing Empathy: The Role of a “Control Room” in an E-Learning Environment. *Interactive Technology and Smart Education*, *10*(1), 31–48. https://doi.org/10.1108/17415651311326437

Henry, E. a, Schlegel, R. J., Talley, A. E., Molix, L. A., & Bettencourt, B. A. (2010). The feasibility and effectiveness of expressive writing for rural and urban breast cancer survivors. *Oncology Nursing Forum*, *37*(6), 749–757. https://doi.org/10.1188/10.ONF.749-757

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

Manier, D., & Olivares, A. (2005). Who benefits from expressive writing? Moderator variables affecting outcomes of emotional disclosure interventions. *Counseling & Clinical Psychology Journal*, *2*(1), 15–28.

Park, C. L., & Blumberg, C. J. (2002). Disclosing trauma through writing: Testing the meaning-making hypothesis. *Cognitive Therapy and Research*, *26*(5), 597–616. https://doi.org/10.1023/A:1020353109229

Pennebaker, J. W. (1989). Confession, inhibition, and disease. In *Advances in Experimental Social Psychology* (Vol. 22, pp. 211–244). https://doi.org/10.1016/S0065-2601(08)60309-3

Pennebaker, J. W. (1997). Writing about emotional experiences as a therapeutic process. *Psychological Science*, *8*(3), 162–166. https://doi.org/10.1111/j.1467-9280.1997.tb00403.x

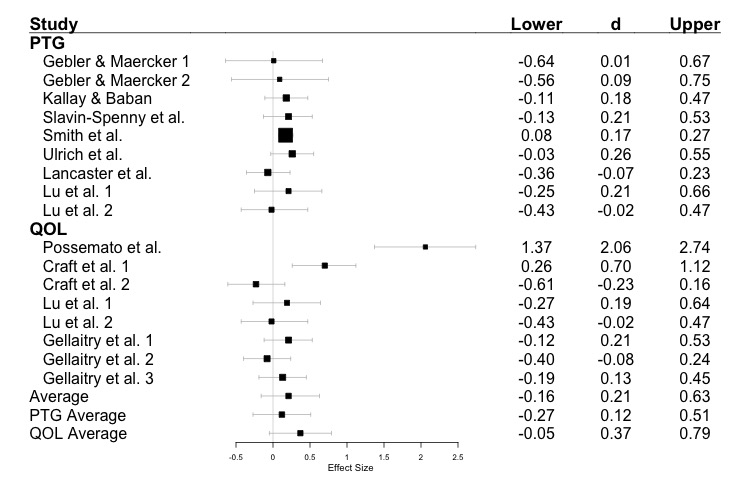
Pennebaker, J. W., & Beall, S. K. (1986). Confronting a traumatic event: Toward an understanding of inhibition and disease. *Journal of Abnormal Psychology*, *95*(3), 274–281. https://doi.org/10.1037//0021-843X.95.3.274

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. https://doi.org/10.1016/j.beth.2006.06.005

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. https://doi.org/10.1016/j.brat.2011.02.001

Smyth, J. M., & Pennebaker, J. W. (2008). Exploring the boundary conditions of expressive writing: In search of the right recipe. *British Journal of Health Psychology*, *13*(1), 1–7. https://doi.org/10.1348/135910707X260117

Yilmaz, M., & Zara, A. (2016). Traumatic loss and posttraumatic growth: the effect of traumatic loss related factors on posttraumatic growth. *Anatolian Journal of Psychiatry*, *17*(1), 5–11. https://doi.org/10.5455/apd.188311



*Figure 1.* Forest plot with average Cohen’s *d* effect sizes and noncentral confidence intervals. Numbers indicate the comparison outlined in Table 1.

Table 1*.*

*Study Characteristics and Results for PTG and QOL Studies*

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Outcome | *N* | *d* | *SE* | *Lower* | *Upper* | *z* | *p* | *Power1* | *Power2* |
| Gebler & Maercker | PTG | 9 | 0.01 | 0.33 | -0.64 | 0.66 | 0.03 | 0.98 | .05 | .06 |
| Gebler & Maercker | PTG | 9 | 0.09 | 0.33 | -0.56 | 0.74 | 0.27 | 0.79 | .06 | .06 |
| Kallay & Baban | PTG | 45 | 0.18 | 0.15 | -0.11 | 0.47 | 1.20 | 0.23 | .22 | .14 |
| Slavin-Spenny | PTG | 36 | 0.21 | 0.17 | -0.12 | 0.54 | 1.25 | 0.21 | .23 | .12 |
| Smith et al. | PTG | 139 | 0.17 | 0.09 | 0.00 | 0.34 | 1.95 | 0.05 | .51 | .33 |
| Ulrich | PTG | 47 | 0.26 | 0.15 | -0.03 | 0.55 | 1.77 | 0.08 | .42 | .14 |
| Lancaster | PTG | 45 | -0.07 | 0.15 | -0.36 | 0.22 | -0.47 | 0.64 | .07 | .14 |
| Lu et al. | PTG | 19 | 0.21 | 0.23 | -0.24 | 0.66 | 0.91 | 0.36 | .14 | .08 |
| Lu et al. | PTG | 19 | -0.02 | 0.23 | -0.47 | 0.43 | -0.09 | 0.93 | .05 | .08 |
| **Possemato et al.** | **QOL** | **48** | **2.06** | **0.19** | **1.68** | **2.44** | **10.66** | **<.01** | **1.00** | **.14** |
| Craft et al. (baseline to one month) | QOL | 26 | 0.70 | 0.21 | 0.29 | 1.11 | 3.38 | <.01 | .93 | .10 |
| Craft et al. (six months) | QOL | 26 | -0.23 | 0.2 | -0.62 | 0.16 | -1.17 | 0.24 | .20 | .10 |
| Lu et al. | QOL | 19 | 0.19 | 0.23 | -0.26 | 0.64 | 0.82 | 0.41 | .12 | .08 |
| Lu et al. | QOL | 19 | -0.02 | 0.23 | -0.47 | 0.43 | -0.09 | 0.93 | .05 | .08 |
| Gellaitry | QOL | 38 | 0.21 | 0.16 | -0.11 | 0.53 | 1.29 | 0.20 | .24 | .12 |
| Gellaitry | QOL | 38 | -0.08 | 0.16 | -0.40 | 0.24 | -0.49 | 0.62 | .08 | .12 |
| Gellaitry | QOL | 38 | 0.13 | 0.16 | -0.19 | 0.45 | 0.80 | 0.42 | .12 | .12 |

*Note*. 1 Power calculated from the study statistics using *d* and *N.* 2 Power calculated from study sample size and overall estimate of *d* = 0.13. Bolded study indicates outlier effect.

Table 2

*Estimates of Effect Size for Fixed and Random Effects Models*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Fixed Effects | | | Random Effects | | |
|  | Effect Size | Lower | Upper | Effect Size | Lower | Upper |
| All Experiments | 0.22 | 0.14 | 0.30 | 0.24 | 0.02 | 0.46 |
| All Experiments  (no outlier) | 0.13 | 0.05 | 0.22 | 0.13 | 0.04 | 0.22 |
| PTG | 0.14 | 0.04 | 0.25 | 0.14 | 0.04 | 0.25 |
| QOL | 0.35 | 0.22 | 0.48 | 0.36 | -0.13 | 0.87 |
| QOL  (no outlier) | 0.12 | -0.02 | 0.26 | 0.12 | -0.09 | 0.33 |