A meta-analysis of the survival-processing advantage in memory

John E. Scofield¹ · Erin M. Buchanan¹ · Bogdan Kostic¹

Published online: 25 July 2017
© Psychonomic Society, Inc. 2017

Abstract The survival-processing advantage occurs when processing words for their survival value improves later performance on a memory test. Due to the interest in this topic, we conducted a meta-analysis to review the literature regarding the survival-processing advantage, in order to estimate a bias-corrected effect size. Traditional meta-analytic methods were used, as well as the test of excess significance, *p*-curve, *p*-uniform, trim and fill, PET-PEESE, and selection models, to reevaluate previous effect sizes while controlling for forms of small-study-size effects. The average effect sizes for survival processing ranged between $\eta_p^2 = .06$ and $.09$ for between-subjects experiments and between $\eta_p^2 = .15$ and $.18$ for within-subjects experiments, after correcting for potential bias and selective reporting. Overall, researchers can expect to find medium to large survival-processing effects, with selective reporting and bias-correcting techniques typically estimating lower effect sizes than traditional meta-analytic techniques.

Keywords Survival processing · Meta-analysis · Memory · Effect size

Evolutionary psychologists view ancestral environments as an integral part of evolutionary psychology. They place importance on this aspect and posit that our brains contain adaptations in the form of modules. These modules help solve adaptive problems that were present in our ancestral environments (Cosmides & Tooby, 1994; Tooby & Cosmides, 2005). Nairne

and Pandeirada (2008) argued that memory is adapted to remember information that is specifically relevant to our evolutionary fitness. That is, our memory systems have evolved around the presence of survival contexts. Studies spanning back to 2007 have shown that processing words for survival value improved later performance on a memory test, and that this memory has significant adaptive value (Nairne, Thompson, & Pandeirada, 2007).

In a typical survival-processing paradigm, participants imagine they are stranded in a grasslands environment without any survival materials and are susceptible to predation and other potential dangers. Participants view and rate a list of words on the basis of how relevant each item would be in that survival scenario. Following the rating procedure, participants receive a memory test and write down as many items as they can recall from the list they previously viewed (Nairne, Pandeirada, & Thompson, 2008). Different encoding instructions can be implemented for comparison, in which participants might imagine themselves in a scenario involving moving to a new country, a neutral processing task such as rating for pleasantness, or other possible scenarios that vary in their degrees of evolutionary relevance. Results typically show that the survival scenario leads to an advantage in memory as compared to other types of encoding procedures. For instance, Renkewitz and Müller (2015) successfully replicated Nairne et al. (2008) as a part of the Reproducibility Project (Nosek et al., 2015). By employing similar methods, Renkewitz and Müller showed that words rated for their survival relevance were recalled at a greater rate than those presented in a vacation scenario. This survival-processing advantage has also been replicated with different comparison conditions, stimulus materials, encoding scenarios, and populations. Survival processing has been shown to persist across different types of designs as well as across recall and recognition tests (Aslan & Bauml, 2012; Bell, Röer, & Buchner, 2013; Kostic,

✉ John E. Scofield
John1551@live.missouristate.edu

¹ Department of Psychology, Missouri State University, 901 S. National Avenue, Springfield, MO 65897, USA

McFarlan, & Cleary, 2012; Nairne & Pandeirada, 2010; Otgaar, Smeets, & van Bergen, 2010; Röer, Bell, & Buchner, 2013; Weinstein, Bugg, & Roediger, 2008).

Nairne et al. (2007) described a functional–evolutionary perspective (also known as an *ultimate explanation*) as their theoretical basis regarding the memorability of survival-relevant items. From this perspective, memory systems have evolved to remember items and information relevant to survival, so processing in terms of the relation to survival improves retention (Burns, Hwang, & Burns, 2011; Nairne et al., 2008). The survival-processing advantage has also been examined from a structuralist or proximate approach, which focuses on the relevant mechanism and explanation of a phenomenon (Burns et al., 2011; Butler, Kang, & Roediger, 2009; Howe & Derbish, 2010; Kang, McDermott, & Cohen, 2008; Kroneisen & Erdfelder, 2011; Soderstrom & McCabe, 2011). Some researchers using this approach have proposed that general principles of memory can explain the survival-processing advantage, without any need to involve fitness relevance (Howe & Otgaar, 2013).

Kostic et al. (2012) and Soderstrom and McCabe (2011) provided evidence that survival-processing effects do not uniformly require ancestral survival scenarios as a prerequisite, by extending this effect across different settings that were not relevant to evolution. Kroneisen and Erdfelder (2011) hold the position that survival processing can be traced back to and explained by richness-of-encoding factors. The richness of encoding of survival conditions allows for elaborate and distinctive encoding more than control tasks or scenarios do, which can explain the increase in retention. Consistent with the richness of encoding, participants generate more uses for items in survival scenarios. When the possibility to spontaneously generate ideas is limited in differing contexts, the survival-processing advantage disappears (Kroneisen & Erdfelder, 2011; Kroneisen, Erdfelder, & Buchner, 2013; Röer et al., 2013).

The role of emotions in survival processing remains uncertain. Kang et al. (2008) compared survival scenarios to a bank heist scenario to examine whether arousal could account for the survival-processing advantage. Word recall was better in the survival than in the bank heist condition, albeit both scenarios received similar excitement/emotional ratings. Smeets, Otgaar, Raymaekers, Peters, and Merckelbach (2012) proposed that survival-processing scenarios could possibly induce stress, which could improve memory from stress hormone effects. However, the effects of stress on memory were found to be independent of the survival-processing effect. Soderstrom and McCabe (2011) also examined the influence of emotions by having participants imagine having to protect themselves against zombies. Even though arousal and valence ratings were higher in this zombie condition than in the original grasslands scenario, they were not significant covariates and could not account for the recall differences.

Thinking about death and the effects of mortality salience on memory has been proposed to explain the survival-processing advantage. Burns, Hart, Kramer, and Burns (2014) considered the “dying-to-remember” (DTR) effect, in which participants are placed in a mortality-salient state to test whether the DTR effect is related to survival processing. Burns et al. found both an association with the DTR effect and an increase in item-specific processing. These results posit that the mechanisms at hand for survival-processing and DTR effects share a relationship and overlap. However, Bell et al. (2013) found that the survival-processing advantage was not due to negativity or mortality salience.

False memory is another area investigated in relation to the survival-processing advantage. Survival-related words were found to be more susceptible than neutral words to false memory effects, and processing survival-related words in terms of their relevance to survival increased susceptibility to false memory (Howe & Derbish, 2010). Even though survival-related information that included false memories resulted in worse, not better, memory, it still could have adaptive significance; for instance, false memories make better primes for problem solving (Garner & Howe, 2014; Howe & Derbish, 2010).

Overall, although there have been many successful replications of the survival-processing advantage in memory, the research on survival processing has provided inconclusive evidence regarding the explanation behind the phenomenon. Work has examined both the ultimate and proximate explanations behind survival processing. For instance, Burns et al. (2011) proposed that survival processing fosters both relational and item-specific processing for encoding, superior to that from control strategies. Bröder, Krüger, and Schütte (2011), however, reported that functional analysis in the case of survival processing is inconclusive. An adaptive memory, Bröder et al. pointed out, should include not only item memory, but aspects including source memory that would also be enhanced by survival value; however, they did not find any effects of adaptive source memory. Nairne and Pandeirada (2010) concurred and did not find survival advantages for source memory. The survival-processing advantage also vanished under dual-task conditions, implying that survival processing may not be prioritized in dual-task contexts (Kroneisen, Rummel, & Erdfelder, 2014). McBride, Thomas, and Zimmerman (2013) pointed out that if memory serves adaptive needs, it might have evolved in an older form of memory. However, McBride et al. found no support for a survival advantage in implicit tests. In sum, although the functional–evolutionary approach of survival processing has certainly received support, work is still ongoing.

Meta-analyses provide us with an alternative to the normal, narrative discussions of research studies, which is useful when attempting to make sense of quickly expanding research (Glass, 1976; Wolf, 1986). The purposes of this meta-

analysis were to review the literature regarding the survival-processing advantage and aggregate the available data to show an averaged effect of the survival-processing advantage. Effect sizes differ for various reasons, potentially as a result of different random sampling (Francis, 2012). Therefore, effect sizes are understood better pooled together across studies. In this article we addressed the aggregation of available studies by performing a literature search, assessing the homogeneity of effect size parameters, pooling the effect sizes, testing and correcting for potential forms of bias, and determining whether any characteristics of these studies were systematically related to the effect size (Borenstein, 2009; Dunlap, Cortina, Vaslow, & Burke, 1996; Hedges, 1982; Hedges & Olkin, 1985; Morris & DeShon, 2002).

Method

Data collection

All materials, including datasets and R code are available online at <http://osf.io/6sd8e>. A literature search was conducted to identify articles related to a survival-processing advantage utilizing search engines and a manual reference check. Search terms for the literature search included memory, adaptive memory, survival processing, and evolution. We first conducted searches on PsycINFO and Google Scholar to locate articles on the survival-processing advantage, and then manually retrieved sources cited within each individual study. Other mediums, such as psychfiledrawer.org and PsyArXiv preprints were also used to search for articles. The dependent variable of interest was word recall from differing types of encoding strategies. Overall, these searches returned 49 relevant studies comprised of 113 experiments, which included seven unpublished dissertations and theses. Ninety experiments were acceptable for inclusion in this analysis. Twenty-three experiments were excluded from this analysis for insufficient quantitative information or differing measures of interest. A table of the specific experiments with exclusion explanations can be found at <http://osf.io/6sd8e>. Of the initially included experiments, 54.4% (49) utilized between-subjects designs, and 45.6% (41) utilized within-subjects designs. All experiments included in the analysis were published between 2007 and 2015.

Effect size, variance, and confidence interval calculations

Test statistics were extracted from all experiments. Only main effects of survival-processing condition were considered. No test statistics were pulled from interactions or main effects unrelated to processing condition. A table of experiment characteristics, including each experiment's author, year of publication, research design, effect size, and sample size is available online at <http://osf.io/6sd8e>. The majority of experiments

(90%, 81) used analyses of variance (ANOVAs) to assess their hypotheses, and therefore, we used partial eta-squared as our measure of effect size. Only one effect size was used from each specific experiment; thus, all effect sizes were independent. If a study contained multiple experiments, only the main effect of survival processing from each experiment was calculated. Regardless of whether effect sizes were reported in the original articles, these values were recalculated from the F ratio and df statistics presented for each experiment using Eq. 1, provided by Cohen (1965):

$$\eta_p^2 = \frac{F \times df_{\text{effect}}}{F \times df_{\text{effect}} + df_{\text{error}}} \quad (1)$$

All effect sizes were treated as partial eta-squareds because, with one-way between-subjects ANOVA designs, full and partial eta values are the same. In one-way within-subjects designs, full eta-squared would traditionally be calculated by using the sum-of-squares total (i.e., sum of squares model + sum of squares subject + sum of squares residual) as the denominator. However, most software packages calculate partial eta-squared by excluding the sum of squares subject variance, and thus we assumed that values reported in the primary studies were partial eta-squared. Eight experiments used t tests, and therefore reported Cohen's d . For experiments utilizing t tests, the Cohen's d values were first converted into correlation coefficients according to Eq. 2, provided by Cooper, Hedges, and Valentine (2009):

$$r = \frac{d}{\sqrt{d^2 + a}}, \quad (2)$$

where a is a correction factor¹ when $n_1 \neq n_2$. This conversion is based on a design with two-independent means. An adequate formula has not currently been derived for transforming Cohen's d values based on within-subjects designs. Although only six of the 90 experiments reported Cohen's d using a within-subjects design, the use of this formula may have caused biased transformations of those effect sizes. The correlation coefficients were then converted into partial eta-squared with Eq. 3:

$$\eta_p^2 = r^2. \quad (3)$$

We followed traditional meta-analytic procedures, detailed in Cooper et al. (2009). The primary effect sizes were weighted in terms of their inverse variances (Sánchez-Meca & Marín-Martínez, 2008). Confidence intervals were also calculated for every meta-analytic effect size reported in the present project. To estimate the sampling variance of partial eta-squared, the primary effect sizes were first converted to raw correlation coefficients. This procedure was advantageous because of the dearth

¹ Alternatively, meta-analyses can also be performed converting all partial eta-squared values into Cohen's d values.

of literature indicating how to estimate variance for eta-type measures, which is a crucial component of the techniques chosen here. Borenstein (2009) advises against analyses performed using correlation coefficients, because the variances are themselves dependent on the correlation, and raw correlations are not normally distributed as are coefficients transformed to Fisher's z scale. To circumvent this problem, correlation coefficients were converted to Fisher's z scale via Eq. 4:

$$z = 0.5 \times \ln\left(\frac{1+r}{1-r}\right). \quad (4)$$

The variance of z could then be estimated per Eq. 5:

$$V_z = \frac{1}{n-3}, \quad (5)$$

where the standard error was the square root of that variance estimation. Traditional meta-analytic procedures were then followed using z . Both fixed- and random-effects models are reported, pooling the primary effect sizes across all experiments (Hedges & Olkin, 1985; Hedges & Vevea, 1998; Marín-Martínez & Sánchez-Meca, 2010; Sánchez-Meca & Marín-Martínez, 2008). However, considering that pooling primary effect sizes across different research designs may not be appropriate, separating experiments according to research design served to make the groups more homogeneous. The primary effect sizes were binned into subgroups (between-subjects or within-subjects designs). The primary effect sizes for the fixed-effects model were weighted in terms of their inverse variances,

$$w_i^{FE} = \frac{1}{V_z}, \quad (6)$$

and then pooled across experiments according to that weight. For the random-effects model, the primary effect sizes were again weighted in terms of their inverse variances, this time using the summation of an experiment's within- and between-study variances,

$$w_i^{RE} = \frac{1}{V_z + \tau^2}, \quad (7)$$

where τ^2 refers to the between-study variance. Between-study variance was estimated using the Paule–Mandel estimator (Langan, Higgins, & Simmonds, 2017; Veroniki et al., 2016). After meta-analytic analyses were performed, the meta-analytic effect size z was converted back into partial eta-squared for presentation, using Eq. 8:

$$\eta_p^2 = \left(\frac{e^{2z}-1}{e^{2z}+1} \right)^2. \quad (8)$$

Confidence intervals for both the primary and meta-analytic effect sizes were calculated using normal-distribution

calculations with the *metafor* package (Viechtbauer, 2010), and all analyses in this article are based on these normal-distribution confidence intervals. However, the literature on effect sizes has indicated that noncentral confidence intervals are potentially more appropriate (Cumming, 2012; Kelley, 2007; Smithson, 2003), and therefore these noncentral F distribution estimates can be found in the supplementary online material for comparison. Forest plots are presented in Figs. 1 and 2 for graphical representation of the primary and meta-analytic effects, with confidence intervals. Forest plots show each study's primary effect size, with each box's size corresponding to the weight of each study and horizontal lines indicating the confidence interval for each individual effect size. Considering the large number of experiments, they were plotted and separated by research design (between vs. within subjects).

Outlier and influential study detection

Before reporting meta-analytic effect sizes, we assessed outliers and influential cases using the *metafor* package, by calculating Studentized deleted residuals, DFBETAS values, and the ratio of generalized variances (COVRATIO; Viechtbauer & Cheung, 2010). Studentized deleted residuals, an outlier identification technique, compares the observed effect size values with those from models that exclude each respective study. Testing for influence, in addition to testing for outliers, is important because the presence of an outlying experiment may not necessarily change, or influence, specific conclusions. An experiment is deemed influential if excluding that experiment yields changes in the fitted model (Viechtbauer & Cheung, 2010). DFBETAS and generalized variance ratio techniques both indicate influential experiments. DFBETAS values indicate the overall change in effect size after excluding each respective experiment from the initial model fitting. DFBETAS values exceeding 1 indicate influential experiments. COVRATIO values smaller than 1 suggest that the exclusion of the i th experiment yields more precision in model coefficient estimates (Viechtbauer & Cheung, 2010). Outliers/influential cases were tested using both a random-effects and a fixed-effects model, and three common outliers/influential cases between the two types of models were identified (a table of the specific identified outliers is available in the online supplemental material).

By definition, however, studies acceptable for inclusion in a meta-analysis may not be considered outliers. Furthermore, Viechtbauer and Cheung (2010) noted that the detection of an outlying/influential study does not automatically merit its deletion. Therefore, a sensitivity analysis was conducted with models that either included or excluded those outlying/influential cases. The sensitivity analysis (available in the online supplementary materials) revealed that the main results and conclusions remained the same whether those cases were included or excluded. All subsequent analyses, knowing that

Fig. 1 This forest plot shows each between-subjects experiment's effect size estimate (listed as TE, with standard error seTE), with boxes according to each experiment's weight and horizontal lines depicting

confidence intervals. W corresponds to the percent weight for each experiment. It is important to note that these estimates are presented on Fisher's z scale and can be converted to partial eta-squared using Eq. 8

the main results/conclusions remained the same, excluded these specific cases in order to reduce the amount of heterogeneity to acceptable levels for bias-correcting techniques. The value of τ^2 was estimated to be 0.02 including the detected cases, whereas excluding those three identified cases yielded a τ^2 of 0.01. The change in τ^2 (52.1%; see Viechtbauer & Cheung, 2010, for details) suggested that the exclusion of these influential cases decreased the estimates of heterogeneity. The exclusion of outlying/influential cases from the present meta-analysis was also advantageous, considering that high heterogeneity estimates can be problematic

for analyses such as p -curve, p -uniform, trim-and-fill methods, and selection models. High-heterogeneity limitations are addressed in the heterogeneity results and in the Discussion section below.

Results

Traditional fixed-effects and random-effects meta-analyses, with outliers/influential cases excluded, revealed overall meta-analytic effect size estimates of $\eta_p^2 = .11$, 95% CI [.10,

Fig. 2 This forest plot shows each within-subjects experiment's effect size estimate (listed as TE, with standard error seTE), with boxes according to each experiment's weight and horizontal lines depicting

confidence intervals. W corresponds to the percent weight for each experiment. It is important to note that these estimates are presented on Fisher's z scale and can be converted to partial eta-squared using Eq. 8

.12], and $\eta_p^2 = .12$, 95% CI [.10, .14], respectively. However, a comparison of the primary effect sizes across different research designs revealed that the experiments using within-subjects designs had higher effect sizes than those using between-subjects designs, Welch-corrected $t(67.68) = -3.46$, $p < .001$, $d = 0.73$, $BF_{10} = 49.02$. Bayes factors were calculated with the *BayesFactor* package (Morey & Rouder, 2015), with a standard Cauchy prior and $r_{scale} = 1$. We created homogeneous subgroups based on the type of research design and fitted meta-analytic models separately. Alternatively, a single model could be fitted using the type of research design as a categorical moderating variable. Both of these options are viable; however, Viechtbauer (2010) recommended fitting models separately due to differences in primary effect size and heterogeneity estimates across levels of a categorical variable, which was the case across the types of research design. Considering the between-subjects experiments, traditional

fixed- and random-effects meta-analyses revealed overall meta-analytic effect size estimates of $\eta_p^2 = .09$, 95% CI [.07, .10], and $\eta_p^2 = .09$, 95% CI [.07, .11], respectively. Fixed and random effects meta-analyses for within-subjects designs showed overall meta-analytic effect size estimates of $\eta_p^2 = .17$, 95% CI [.15, .20] and $\eta_p^2 = .17$, 95% CI [.14, .21]. The forest plot for between-subjects designs is presented in Fig. 1, and that for within-subjects designs is in Fig. 2. Table 1 shows the fixed- and random-effects model estimates for the overall, between-subjects, and within-subjects designs across all results described below.

Homogeneity

The homogeneity of the meta-analytic results was assessed using both the Q statistic with a chi-square distribution ($k - 1$ df), wherein k was the number of studies (Cochran, 1954;

Table 1 Effect size estimates across meta-analytic methods

All Studies	Fixed Effects	Random Effects
Traditional	.11 [.10, .12]	.12 [.10, .14]
<i>P</i> -uniform	.11 [.08, .14]	–
Trim and fill	.07 [.06, .09]	.08 [.06, .10]
PET–PEESE	–	.07 [.04, .11]
Selection models	.11 [.09, .12]	.09 [.06, .13]
Between-Subjects		
Traditional	.09 [.07, .10]	.09 [.07, .11]
<i>P</i> -uniform	.09 [.06, .11]	–
Trim and fill	.07 [.06, .08]	.07 [.05, .09]
PET–PEESE	–	.06 [.03, .10]
Selection models	.08 [.07, .10]	.08 [.05, .11]
Within-Subjects		
Traditional	.17 [.15, .20]	.17 [.14, .21]
<i>P</i> -uniform	.16 [.10, .22]	–
Trim and fill	.18 [.15, .21]	.18 [.14, .22]
PET–PEESE	–	.18 [.10, .28]
Selection models	.17 [.14, .21]	.15 [.09, .22]

This table shows partial eta-squared estimates across experiments along with 95% normal confidence intervals for the various types of research synthesis analyses, with the last four techniques correcting for selective reporting.

Huedo-Medina, Sánchez-Meca, Marín-Martínez, & Botella, 2006), and the I^2 index. The Q statistic is akin to the variance of the effects, as a between-study weighted sum of the squared differences, which indicates the variation in between-study effects. I^2 is similar to the Q statistic, which quantifies the degree of heterogeneity versus chance in a meta-analysis, and therefore both values are reported. I^2 is often considered to measure the inconsistency between studies (Higgins, Thompson, Deeks, & Altman, 2003). However, I^2 estimates can be inexact when a meta-analysis includes a small number of experiments or the sample size within the experiments is small. Wide confidence intervals can be understood as a manifestation of these inexact estimates, which are related to power issues found with the Q test. Therefore, 95% confidence intervals are reported along with every I^2 statistic. The reasons why tests could fail in terms of homogeneity include the use of different measures or designs, or even because the sampling of subjects differed greatly (Huedo-Medina et al., 2006; Wolf, 1986). These values were calculated using the *meta* package.

We rejected the homogeneity assumption for the overall effects, $Q(86) = 175.04$, $p < .001$. The amount of variability among the effect sizes caused by true heterogeneity across experiments was $I^2 = 50.9\%$, 95% CI [37.1%, 61.6%]. We then assessed homogeneity for the between-subjects design experiments, $Q(47) = 75.25$, $p = .01$. According to the I^2 index, heterogeneity was reduced to a low to moderate level when split in this way by research design, $I^2 = 37.5\%$, 95% CI

[11.2%, 56.1%]. For experiments implementing a within-subjects design, homogeneity was again rejected, $Q(38) = 69.67$, $p < .01$. The I^2 index, however, indicated a moderate level of heterogeneity, $I^2 = 45.5\%$, 95% CI [20.4%, 62.6%]. Excluding outlying/influential cases and breaking experiments into homogeneous subgroups helped decrease heterogeneity, thus making the use of bias-correcting techniques appropriate.

Test for excess significance

The test for excess significance (TES; Ioannidis & Trikalinos, 2007) was used to determine the likelihood of rejecting the null hypothesis given a specific number of experiments. The logic of the TES is similar to that of hypothesis testing, in the sense that the null hypothesis in the TES is that experiments were run properly and without bias. The use of the TES is appropriate for individual studies with four or more experiments (Francis, 2012, 2014). Statistical power for each experiment included in the TES analysis was calculated (with the *pwr* package; Champely, 2009) using the meta-analytic effect size estimate appropriate for which research design was implemented. For example, if survival processing was manipulated between subjects in an experiment, power was calculated using the traditional between-subjects meta-analytic effect size, not that primary study's effect size. Francis (2012, 2014) details this procedure, indicating that if the product of all the experiments' power estimates from each study falls below the suggested .10 criterion, this finding indicates more rejections of the null hypothesis than would be expected from the estimated parameters.

Evidence for survival processing is not reliant simply on results from an entire meta-analysis, because these results cover a broad range of experimental manipulations. Therefore, the TES was not implemented across the entire meta-analysis. By concentrating on individual articles, we could identify the specific theoretical claims tested by experiments within a single study. Even potential bias overall does not reveal much regarding singular aspects of survival processing. Therefore, this problem was avoided by applying the TES to individual studies. Considering that the TES was only implemented with sets of experiments containing all significant findings, the p -values from all primary studies were first recalculated using the *ci* function from the *meta* package in R (Schwarzer, Carpenter, & Rücker, 2015). Recalculating p values allowed us to accurately determine which sets of experiments were viable for inclusion in the TES. In all, 82.6% (71) of experiments reported significant results. Four studies (with significant findings across all experiments) out of the 40 included in the meta-analysis were acceptable to be examined via the TES, since they had at least four experiments (Francis, 2012). After power was calculated for each experiment, as described above, the product of all the experiments' power

estimates was calculated. The present TES analysis showed that none of the four studies acceptable for analysis fell below the .10 threshold. Table 2 shows each study and its experiments, the type of design implemented, the power estimates, and the probability of excess significance.

P-curve and *p*-uniform

A *p*-curve analysis examines the *p*-value distributions of statistically significant findings and indicates whether a set of experiments contains evidential value (van Aert, Wicherts, & van Assen, 2016). The underlying logic of *p*-curve compares differences in the distributions of statistically significant *p* values. When no effect exists, given a set of data, *p* values will be uniformly distributed (flat). If a field contains evidential value, the *p*-value distribution will appear right-skewed (Simonsohn, Nelson, & Simmons, 2014). Along with testing for evidential value, *p*-curve analysis also provides statistical power estimates after correcting for selective reporting

(Simonsohn, Simmons, & Nelson, 2015). All *p*-curve analyses were performed using the online application at *p*-curve.com (Simonsohn et al., 2014). *p*-uniform, an alternative to *p*-curve analysis, also examines *p*-value distributions. With *p*-uniform analysis, the population effect size underlying the effect sizes from primary studies is assumed to be the same. This analysis is referred to as “uniform” because the *p*-value distribution for the population effect is considered uniform when tested against the true effect size (van Assen, van Aert, & Wicherts, 2014). The use of *p*-uniform also includes a formal test for publication bias and can estimate corrected meta-analytic effect sizes. Simonsohn et al. (2014) initially proposed effect size estimation via *p*-curve. However, through personal email communication with one of the authors of Simonsohn et al., we found that effect size estimation via *p*-curve was deemed inappropriate (U. Simonsohn, personal communication, October 25, 2016). Thus, effect size estimation with *p*-uniform served as a valuable alternative. The *p*-uniform analyses were performed using the *puniform* package (van Aert, 2017).

The results from the *p*-curve analysis showed that overall the experiments do contain evidential value, $z = -15.05$, $p < .001$ (full *p*-curve with $ps < .05$), and $z = -15.90$, $p < .001$ (half *p*-curve with $ps < .025$). Figure 3 shows the observed *p*-curve relative to both a uniform distribution and a distribution with 33% power. Furthermore, the evidential value inferred from *p*-curve does not indicate evidential inadequacy, $z = 8.70$, $p > .999$. The results from the *p*-uniform analysis revealed no indication of publication bias, $z = 0.17$, $p = .43$.

These results are mirrored when split by between- and within-subjects designs. Both types of research designs do contain evidential value: between-subjects full *p*-curve $z = -8.74$, $p < .001$, and half *p*-curve $z = -8.65$, $p < .001$; within-subjects full *p*-curve $z = -12.52$, $p < .001$, and half *p*-curve $z = -13.88$, $p < .001$. Figures 4 and 5 portray the observed *p*-curves for between-subjects and within-subjects designs, respectively. Neither test indicated evidential inadequacy: between-subjects $z = 4.35$, $p > .999$, within-subjects $z = 7.93$, $p > .999$. Finally, no indication of publication bias was found using *p*-uniform: between-subjects $z = -0.05$, $p = .52$; within-subjects $z = 0.59$, $p = .28$. The *p*-uniform fixed-effect estimates are shown in Table 1. Figures plotting expected conditional *p* values against observed conditional *p* values from all *p*-uniform analyses are available in the supplementary materials online.

Trim and fill

The trim-and-fill method is based on the relationship between the primary effect size estimates and their corresponding standard errors, and how that relationship may change in the presence of small-study effects such as publication bias (Carter & McCullough, 2014). Funnel plots graphically display the

Table 2 Test for excess significance (TES) results

Author	Design	Power	
Kostic et al. (2012), Exp. 1a	WI	.83	
Kostic et al. (2012), Exp. 1b	BN	.81	
Kostic et al. (2012), Exp. 2a	WI	.99	
Kostic et al. (2012), Exp. 2b	WI	.99	
		.66	P_{TES}
Naime & Pandeirada (2011), Exp. 1a	WI	.57	
Naime & Pandeirada (2011), Exp. 1b	WI	.57	
Naime & Pandeirada (2011), Exp. 2	BN	.89	
Naime & Pandeirada (2011), Exp. 3	WI	.99	
Naime & Pandeirada (2011), Exp. 4	BN	.98	
		.29	P_{TES}
Naime et al. (2007), Exp. 1	BN	.93	
Naime et al. (2007), Exp. 2	WI	.71	
Naime et al. (2007), Exp. 3	WI	.73	
Naime et al. (2007), Exp. 4	WI	.83	
		.40	P_{TES}
Otgaar & Smeets (2010), Exp. 1	BN	.60	
Otgaar & Smeets (2010), Exp. 2	BN	.77	
Otgaar & Smeets (2010), Exp. 2	BN	.62	
Otgaar & Smeets (2010), Exp. 3	BN	.47	
		.14	P_{TES}

This table shows each study (listed by experiment) included in the TES analysis. Each experiment is listed with the authors, type of research design (BN, between; WI, within), and statistical power calculated on the basis of the pooled meta-analytic effect size estimate, according to its corresponding research design. P_{TES} refers to the probability of the observed, or better, significance level, with a score lower than .10 indicating excessive significance given the study parameters. The *p* values for all experiments are available at <http://osf.io/6sd8e>

Fig. 3 This graph depicts the observed p -curve for all experiments, compared to both a uniform distribution and a 33%-power distribution. The observed p -curve includes 73 statistically significant ($p < .05$) results,

of which 57 have $p < .025$. An additional 14 results were entered but are excluded from the p -curve because they had $p > .05$

spread of primary effect sizes along the x -axis, with standard error or precision (inverse of the standard error) along the y -axis. Funnel plots can indicate potential asymmetry, wherein a

lack of data points in the lower center area of the plot indicates studies with nonsignificant findings and corresponding small sample sizes (i.e., a funnel plot asymmetry; van Assen et al.,

Fig. 4 This graph shows the observed p -curve for the between-subjects experiments subgroup, compared to both a uniform distribution and a 33%-power distribution. The observed p -curve includes 36 statistically

significant ($p < .05$) results, of which 29 have $p < .025$. An additional 12 results were entered but are excluded from the p -curve because they had $p > .05$

Fig. 5 This graph shows the observed p -curve for the within-subjects experiments subgroup, compared to both a uniform distribution and a 33%-power distribution. The observed p -curve includes 37 statistically

significant ($p < .05$) results, of which 28 have $p < .025$. An additional two results were entered but are excluded from the p -curve because they had $p > .05$

2014). The trim-and-fill method first “trims” a given funnel plot until the data points within the plot are symmetrical. Next, the number of missing studies is estimated and imputed, or “filled,” all while maintaining symmetry within the funnel plot (Duval & Tweedie, 2000). The trim-and-fill method also estimates corrected meta-analytic effect sizes after the data points are imputed. Trim-and-fill analyses were performed using the *meta* package in R. Figure 6 indicates funnel plots for the overall, between-subjects, and within-subjects data, and Fig. 7 includes the trim-and-fill estimations. The trim-and-fill method had $k = 110$ with 23 added studies for the overall set of experiments (between: $k = 55$, with 7 added; within: $k = 40$, with 1 added). Meta-analytic effect size estimates using the trim-and-fill method for between-subjects designs were slightly lower than those from traditional meta-analytic methods. The meta-analytic effect size estimates for within-subjects designs, in contrast, slightly increased as compared to traditional methods, as can be seen in Table 1.

PET–PEESE

Egger’s regression test is a weighted least squares regression model that evaluates the relationship between the standard error and primary effect size estimates (Egger, Smith, Schneider, & Minder, 1997). Egger’s regression test, used as a test for publication bias, is related to the trim-and-fill method, based on measuring the funnel plot asymmetry (Carter &

McCullough, 2014). Stanley (2005) suggested that, along with testing for funnel plot asymmetry (a significant slope coefficient), the intercept yields an effect size estimate that is void of publication bias. These extensions of Egger’s regression test are referred to as the precision effect test (PET) and the precision effect estimate with standard error (PEESE). PET is more accurate when the true effect size estimates are zero, whereas PEESE is more accurate when the effect size estimates are nonzero. Stanley and Doucouliagos (2013) discussed using PET–PEESE contingent upon whether estimates are zero or nonzero. If $b_0 = 0$, then inferences should be drawn using PET, and if $b_0 \neq 0$, inferences should be drawn using PEESE. PET–PEESE analyses were performed using the *lm* function in R. The results from PET–PEESE indicated a nonzero effect; hence, inferences were drawn using PEESE. The results from PEESE indicated significant funnel plot asymmetry, $b = 5.60$, $t(85) = 2.14$, $p = .01$, $R^2 = .07$. However, this overall effect was not found for either between-subjects studies, $b = 5.44$, $t(46) = 1.60$, $p = .12$, $R^2 = .05$, or within-subjects studies, $b = -0.65$, $t(37) = -0.20$, $p = .84$, $R^2 < .01$, alone. The intercept estimates of meta-analytic effect sizes from this analysis can be seen in Table 1.

Selection models

Finally, selection models offer meta-analytic effect size estimation in the presence of selective reporting. Vevea and

Fig. 6 These funnel plots show effect sizes along the x -axes and their corresponding standard errors along the y -axes for all experiments as well as for the two types of research designs. Funnel plot asymmetry is noted

by a lower frequency of data points at the lower center of the plot. The gray area indicates the region of statistical nonsignificance

Hedges (1995) provided selection models that use the method of maximum likelihood estimation. This selection model offers a formal test for publication bias as well as adjusted effect

size estimations. Selection models were analyzed using the *weightr* package (Coburn & Vevea, 2016). For the overall set of experiments, selection models estimated $\eta_p^2 = .11$,

Fig. 7 These funnel plots show effect sizes along the x -axes and their corresponding standard errors along the y -axes for all experiments as well as for the two types of research designs. These three funnel plots depict

results post-trim-and-fill analysis, after the plots have been trimmed and subsequently imputed. Hollow data points indicate experiments that have been “filled in” or added

95% CI [.09, .12], for fixed effects, with a likelihood ratio test $\chi^2(1) = 1.86$, $p = .17$, indicating no significant difference between selection models and traditional meta-analysis. For random effects, $\eta_p^2 = .09$, 95% CI [.06, .13], with a likelihood ratio test $\chi^2(1) = 7.73$, $p = .01$, indicating that the adjusted model yielded more accurate estimates than traditional meta-analysis (Vevea & Hedges, 1995). However, after separating by design types, we did not find a difference between selection models and traditional estimates of effects for either fixed effects [between: $\chi^2(1) = 0.40$, $p = .52$; within: $\chi^2(1) = 0.02$, $p = .88$] or random effects [between: $\chi^2(1) = 1.93$, $p = .16$; within: $\chi^2(1) = 1.88$, $p = .17$]. The selection model estimates can be seen in Table 1.

Discussion

The survival-processing advantage demonstrates that processing words for their survival value improves memory performance, with a functional approach explaining that this survival processing has an adaptive basis (Nairne et al., 2007). This phenomenon has been replicated in multiple settings and variations (Aslan & Bäuml, 2012; Bell et al., 2013; Kostic et al., 2012; Nairne & Pandeirada, 2010; Nairne et al., 2008; Otgaar et al., 2010; Röer et al., 2013; Weinstein et al., 2008). With respect to effect sizes, we found significant differences between the types of research designs used, with within-subjects experiments yielding higher primary effect sizes than between-subjects experiments. A possible explanation for this result could be that within-subjects designs have increased statistical power as compared to between-subjects designs due to a reduction in the variance across participants (Cohen, 1988), as well as a reduction in the denominator of partial eta-squared because subject variance is excluded. Sample size did appear to be a significant predictor of effect size, $b = -0.001$, $t(88) = -2.04$, $p = .04$, multiple $R^2 = .05$. However, the Bayes factor returned negligible evidence, $BF_{10} = 1.35$. This Bayes factor, calculated using the *BayesFactor* package, used $r_{scale} = \sqrt{2}/4$.

Bias

The results from the PET–PEESE analysis indicated significant funnel plot asymmetry when considering all experiments. This result could indicate publication bias, in which academic journals have the propensity only to publish significant findings. Alternatively, significant funnel plot asymmetry could also stem from authors not submitting articles with nonsignificant findings (Coursol & Wagner, 1986). Often results that fail to meet the acceptable significance threshold do not proceed in the publication process, which leads to what is known as the “file-drawer problem.” With regard to this evidence of

funnel plot asymmetry (which is a type of small-study bias), what does bias mean? Statistically speaking, bias can imply that the frequency of producing a significant result is systematically overestimated (Francis, 2014). Biases could include publication bias, selective analysis, outcome bias, or fabrication bias (Ioannidis & Trikalinos, 2007). Researchers could also be introducing bias into their studies by only reporting successful studies from a potentially larger set containing unsuccessful experiments, as part of the file-drawer problem.

This result raises questions about relying on the results and conclusions of studies containing biases. It is important to note that although funnel plot asymmetry can indicate publication bias, it cannot rule out other potential small-study effects. John, Loewenstein, and Prelec (2012) and Simmons, Nelson, and Simonsohn (2011) have both discussed potential reasons for bias in general, which can include publication bias and/or questionable research practices (e.g., falsifying data, failing to report all dependent measures used, data-peeking, or excluding data after the results are known). However, the results here are difficult to interpret, given that separate analyses did not indicate bias, but did identify that between-subjects designs are likely the culprit behind a significant overall indication of bias. The estimate of the PEESE analysis for between-subjects experiments was greater than five points (i.e., $b = 5.44$), whereas the estimate of the PEESE analysis for within-subjects experiments was very close to zero ($b = -0.65$). Therefore, the larger intercept from between-subjects experiments likely increased the overall intercept ($b = 5.60$) when considering all experiments, and the combined sample size increased the power for the test to detect a difference from zero.

The results from p -uniform, in contrast, did not show evidence of publication bias when considering all experiments. The results from the TES also did not show any evidence of excess significance within the individual studies. However, only four studies with all significant results were analyzed via the TES. No signs of publication bias or funnel plot asymmetry appeared within the subgroup of either the between- or the within-subjects experiments. For all experiments and the between-subjects subgroup, meta-analytic estimates from the techniques typically used returned similar or lower estimates than traditional fixed- and random-effects models. Some techniques used on the within-subjects subgroup returned lower estimates (p -uniform and random-effects selection models), whereas other techniques returned estimates higher than the traditional meta-analytic estimates (trim and fill and PET–PEESE).

Effect size

Although meta-analytic effect size estimates were presented considering all experiments, pooling effect sizes across different research designs was not a suitable choice for this analysis.

Therefore, we recommend caution in interpreting the averaged effect sizes across all experiments. The estimates from both subgroups would be more appropriate in terms of inferences and interpretations, and the estimates across all experiments best serve as a comparison for those of the two subgroups, to note any changes in estimates. Table 1 shows a range of meta-analytic effect size estimates from the different analyses. Inzlicht, Gervais, and Berkman (2015) investigated different types of meta-analytic bias correction techniques and found that no single technique was superior across a range of conditions. Therefore a range of effect sizes should be considered from different types of meta-analytic techniques, especially those that control for selective reporting or bias. For between-subjects experiments, a researcher can expect to find effect sizes for survival processing ranging anywhere between .06 and .09. For within-subjects experiments, researchers can expect to find effect sizes for survival processing ranging between .15 and .18.

Limitations

Limitations of this meta-analysis include biased meta-analytic results, considering that published research is biased in favor of significant findings (Glass, McGaw, & Smith, 1981; Wolf, 1986). If this meta-analysis included all studies that were left unpublished (although a few were located), the averaged effect size potentially would be lower, as we see with some meta-analytic techniques that correct for selective reporting. Okada (2013) posited that eta-squared is not necessarily the best estimator to use, since it itself is slightly biased, especially in the case of small sample sizes. Even after pooling the primary effect sizes by inverse variance and controlling for experiments with small sample sizes, biased eta-squared estimates still could have inflated the meta-analytic estimates in the present project. Because eta-squared is known to have a positive bias, alternatives for effect size estimation include epsilon-squared and omega-squared (Okada & Hoshino, 2016), if one could obtain the appropriate statistics. The results from the present analysis might change slightly, depending on whether normal distributions or noncentral estimates are used. However, confidence interval estimation would likely not change the conclusions of the analysis, considering that the noncentral intervals were still above, and usually did not include, zero. The conclusions from the present meta-analysis also may be difficult to interpret because of aggregated data that included different measuring techniques, definitions of the variables, and subject sampling (Glass et al., 1981; Wolf, 1986).

Caution in interpretation is suggested if results include both well and poorly designed studies. This possibility, however, does not lead us to doubt the conclusions of the present analysis. Experiments with differing measures of interest and insufficient quantitative information were excluded from the

present analysis. All experiments were screened for common outliers and influential cases. Effect size estimates may also be more accurate when considering more homogeneous subgroups. Limitations of the TES analysis include that at least four experiments are required for inclusion, limiting the number of studies viable for this analysis.

Common limitations for *p*-curve, *p*-uniform, trim and fill, PET-PEESE, funnel plots, and selection models include that these techniques are not appropriate to apply in cases in which heterogeneity is high ($I^2 > 50\%$), because they may overestimate meta-analytic effect sizes (van Aert et al., 2016). Heterogeneity in the present analyses was at an acceptable level. However, *p*-uniform can overestimate meta-analytic effect size estimates in the presence of between-study heterogeneity. This heterogeneity could explain why the *p*-uniform estimates were similar to the estimates from traditional meta-analysis and why no publication bias was detected via *p*-uniform. Additionally, with such tests as PET-PEESE we could only test whether publication bias exists within the present set of experiments; this method does not elucidate or distinguish between specific theoretical claims from the experiments (e.g., whether memory systems reflect adaptations specifically for survival processing). Researchers should also be reluctant to interpret effect size estimates in the presence of *p*-hacking (Simonsohn et al., 2014; van Aert et al., 2016). The selection models from Vevea and Hedges (1995) also have a limitation, with potentially inaccurate estimates resulting when the number of studies is less than 100 (Field & Gillett, 2010; van Assen et al., 2014). We therefore recommend caution in interpreting the selection model results. There has been some debate regarding whether the trim-and-fill method should still be implemented in research synthesis. Some consensus points to the conclusion that trim and fill does not yield more valid estimates than any other technique (Carter & McCullough, 2014). Trim and fill has also been reported to undercorrect for publication bias and occasionally to yield inaccurate confidence intervals (Terrin, Schmid, Lau, & Olkin, 2003). The use of funnel plots, which the trim-and-fill method is dependent on, also may be inappropriate with heterogeneous samples or a strong correlation between effect size and sample size (Simonsohn, 2017). Although we do suggest caution in relying solely on funnel plots for meta-analytic inference, we used homogeneous subgroups, and sample size was not a strong predictor of effect size.

The path forward

The initial analyses in the present meta-analysis appeared to reflect potential bias concerning the survival-processing advantage. However, when we separated studies on the basis of type of research design (i.e., between or within subjects), potential bias or small-study-size effects were mitigated, creating a more positive outlook on survival processing.

Van Elk et al. (2015) pointed out that conclusions from meta-analyses are often limited, due to methodological shortcomings. Bias correction techniques often can be inconsistent across a range of conditions, requiring ranges of effect size interpretations instead of the use of a single technique (Inzlicht et al., 2015). Van Elk et al. suggested that a better solution to establish the reliability of an effect (or to potentially decrease reliability) is to focus on large-scale preregistered replications of various phenomena. By implementing preregistered replications, experimenter and publication biases can be eliminated. Suggestions for future research to avoid potential bias and low power involve focusing on confidence intervals and meta-analysis rather than hypothesis testing. Increased sample size and *a priori* study planning can also increase the power of a study. The adoption of Bayesian data analysis methods can be an alternative option to implement (Kruschke, 2010; Nosek et al., 2015; Wagenmakers, 2007). Although Bayesian techniques cannot disclose whether or not an effect exists in the presence of small sample sizes, it does allow for greater flexibility in sampling plans, as well as for not being as reliant on effect size estimation with regard to a priori power analyses (Schönbrodt, Wagenmakers, Zehetleitner, & Perugini, 2017). Finally, meta-analyses like this one, even with their limitations, allow one to examine the merit of research in an area (i.e., nonzero effect sizes) and to plan future studies with better-informed power analyses.

References

- *References marked with an asterisk indicate studies included in the meta-analysis that are not cited in-text.
- *Abel, M., & Bäuml, K. H. T. (2013). Adaptive memory: The influence of sleep and wake delay on the survival-processing effect. *Journal of Cognitive Psychology*, 25, 917–924. doi:10.1080/20445911.2013.825621
- Aslan, A., & Bäuml, K. H. T. (2012). Adaptive memory: Young children show enhanced retention of fitness-related information. *Cognition*, 122, 118–122. doi: 10.1016/j.cognition.2011.10.001
- Bell, R., Röer, J. P., & Buchner, A. (2013). Adaptive memory: The survival-processing memory advantage is not due to negativity or mortality salience. *Memory & Cognition*, 41, 490–502. doi: 10.3758/s13421-012-0290-5
- Bröder, A., Krüger, N., & Schütte, S. (2011). The survival processing memory effect should generalise to source memory, but it doesn't. *Psychology*, 2, 896–901. doi: 10.4236/psych.2011.29135
- *Burns, D. J., Hart, J., Griffith, S. E., & Burns, A. D. (2013). Adaptive memory: The survival scenario enhances item-specific processing relative to a moving scenario. *Memory*, 21, 695–706. doi:10.1080/09658211.2012.752506
- Burns, D. J., Hwang, A. J., & Burns, S. A., (2011). Adaptive memory: Determining the proximate mechanisms responsible for the memorial advantages of survival processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 206–218. doi: 10.1037/a0021325
- Butler, A. C., Kang, S. H., & Roediger, H. L., III. (2009). Congruity effects between materials and processing tasks in the survival processing paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 1477–1486. doi: 10.1037/a0017024
- *Caldwell, J. (2010). *Survival and self-descriptive processing of abstract and concrete nouns* (Unpublished doctoral dissertation). University of Mississippi, Oxford, MS, United States.
- *Ceo, D. A. (2008). *Memory for survival processing of hierarchical categories* (Unpublished doctoral dissertation). Purdue University, West Lafayette, IN, United States.
- *Claxton, A. (2015). *Evolution, memory processes, and the survival processing benefit to memory: An examination of the unpredictability hypothesis* (Unpublished master's thesis). Colorado State University, Fort Collins, CO, United States.
- *Colyn, L. A. (2014). *Planning and the survival processing effect: An examination of the proximate mechanisms* (Unpublished doctoral dissertation). Bowling Green State University, Bowling Green, OH, United States.
- *Fiacconi, C. M., Dekraker, J., & Köhler, S. (2015). Psychophysiological evidence for the role of emotion in adaptive memory. *Journal of Experimental Psychology: General*, 144, 925–933. doi:10.1037/xge0000097
- Garner, S. R., & Howe, M. L. (2014). False memories from survival processing make better primes for problem-solving. *Memory*, 22, 9–18. doi: 10.1080/09658211.2012.759975
- *Giudice, N. D. (2016). *The adaptive effect: Exploring need for cognition and survival processing* (Unpublished master's thesis). University of North Florida, Jacksonville, FL, United States.
- Howe, M. L., & Derbish, M. H. (2010). On the susceptibility of adaptive memory to false memory illusions. *Cognition*, 115, 252–267. doi: 10.1016/j.cognition.2009.12.016
- Kang, S. H., McDermott, K. B., & Cohen, S. M. (2008). The mnemonic advantage of processing fitness-relevant information. *Memory & Cognition*, 36, 1151–1156. doi: 10.3758/MC.36.6.1151
- *Klein, S. B. (2012). A role for self-referential processing in tasks requiring participants to imagine survival on the savannah. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 1234–1242. doi:10.1037/a0027636
- *Klein, S. B. (2013). Does optimal recall performance in the adaptive memory paradigm require the encoding context to encourage thoughts about the environment of evolutionary adaptation? *Memory & Cognition*, 41, 49–59. doi:10.3758/s13421-012-0239-8
- *Klein, S. B., Robertson, T. E., & Delton, A. W. (2011). The future-orientation of memory: Planning as a key component mediating the high levels of recall found with survival processing. *Memory*, 19, 121–139. doi:10.1080/09658211.2010.537827
- Kostic, B., McFarlan, C. C., & Cleary, A. M. (2012). Extensions of the survival advantage in memory: examining the role of ancestral context and implied social isolation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 1091–1098. doi: 10.1037/a0026974
- Kroneisen, M., & Erdfelder, E. (2011). On the plasticity of the survival processing effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 1553–1562. doi: 10.1037/a0024493
- Kroneisen, M., Erdfelder, E., & Buchner, A. (2013). The proximate memory mechanism underlying the survival-processing effect: Richness of encoding or interactive imagery? *Memory*, 21, 494–502. doi: 10.1080/09658211.2012.741603
- Kroneisen, M., Rummel, J., & Erdfelder, E. (2014). Working memory load eliminates the survival processing effect. *Memory*, 22, 92–102. doi: 10.1080/09658211.2013.815217
- Nairne, J. S., & Pandeirada, J. N. (2008). Adaptive memory: Is survival processing special? *Journal of Memory and Language*, 59, 377–385. doi: 10.1016/j.jml.2008.06.001

- Nairne, J. S., & Pandeirada, J. N. (2010). Adaptive memory: Ancestral priorities and the mnemonic value of survival processing. *Cognitive Psychology*, 61, 1–22. doi: [10.1016/j.cogpsych.2010.01.005](https://doi.org/10.1016/j.cogpsych.2010.01.005)
- Nairne, J. S., & Pandeirada, J. N. (2011). Congruity effects in the survival processing paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 539–549. doi: [10.1037/a0021960](https://doi.org/10.1037/a0021960)
- *Nairne, J. S., Pandeirada, J. N., Gregory, K. J., & VanArsdall, J. E. (2009). Adaptive memory: Fitness-relevance and the hunter-gatherer mind. *Psychological Science*, 20, 740–746. doi: [10.1111/j.1467-9280.2009.02356.x](https://doi.org/10.1111/j.1467-9280.2009.02356.x)
- Nairne, J. S., Pandeirada, J. N., & Thompson, S. R. (2008). Adaptive memory: The comparative value of survival processing. *Psychological Science*, 19, 176–180. doi: [10.1111/j.1467-9280.2008.02064.x](https://doi.org/10.1111/j.1467-9280.2008.02064.x)
- Nairne, J. S., Thompson, S. R., & Pandeirada, J. N. (2007). Adaptive memory: Survival processing enhances retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 263–273. doi: [10.1037/0278-7393.33.2.263](https://doi.org/10.1037/0278-7393.33.2.263)
- *Nairne, J. S., VanArsdall, J. E., Pandeirada, J. N., & Blunt, J. R. (2012). Adaptive memory: enhanced location memory after survival processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 495–501. doi: [10.1037/a0025728](https://doi.org/10.1037/a0025728)
- Otgaar, H., & Smeets, T. (2010). Adaptive memory: Survival processing increases both true and false memory in adults and children. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36, 1010–1016. doi: [10.1037/a0019402](https://doi.org/10.1037/a0019402)
- Otgaar, H., Smeets, T., & van Bergen, S. (2010). Picturing survival memories: Enhanced memory after fitness-relevant processing occurs for verbal and visual stimuli. *Memory & Cognition*, 38, 23–28. doi: [10.3758/MC.38.1.23](https://doi.org/10.3758/MC.38.1.23)
- *Palmore, C. C., Garcia, A. D., Bacon, L. P., Johnson, C. A., & Kelemen, W. L. (2011). Congruity influences memory and judgments of learning during survival processing. *Psychonomic Bulletin & Review*, 19, 119–125. doi: [10.3758/s13423-011-0186-6](https://doi.org/10.3758/s13423-011-0186-6)
- *Raymaekers, L. H., Otgaar, H., & Smeets, T. (2014). The longevity of adaptive memory: Evidence for mnemonic advantages of survival processing 24 and 48 hours later. *Memory*, 22, 19–25. doi: [10.1080/09658211.2013.791321](https://doi.org/10.1080/09658211.2013.791321)
- Renkewitz, F., & Müller, S. M. (2015). Replication of Nairne, Pandeirada, & Thompson (2008, PS, Study 2). Retrieved from <http://osf.io/jhkpe>
- Röer, J. P., Bell, R., & Buchner, A. (2013). Is the survival-processing memory advantage due to richness of encoding? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39, 1294–1302. doi: [10.1037/a0031214](https://doi.org/10.1037/a0031214)
- Smeets, T., Otgaar, H., Raymaekers, L., Peters, M. J., & Merckelbach, H. (2012). Survival processing in times of stress. *Psychonomic Bulletin & Review*, 19, 113–118. doi: [10.3758/s13423-011-0180-z](https://doi.org/10.3758/s13423-011-0180-z)
- *Stillman, C. M., Coane, J. H., Profaci, C. P., Howard, J. H., Jr., & Howard, D. V. (2014). The effects of healthy aging on the mnemonic benefit of survival processing. *Memory & Cognition*, 42, 175–185. doi: [10.3758/s13421-013-0353-2](https://doi.org/10.3758/s13421-013-0353-2)
- Weinstein, Y., Bugg, J. M., & Roediger, H. L., III. (2008). Can the survival recall advantage be explained by basic memory processes? *Memory & Cognition*, 36, 913–919. doi: [10.3758/MC.36.5.913](https://doi.org/10.3758/MC.36.5.913)
- Borenstein, M. (2009). Effect sizes for continuous data. In H. Cooper, L. V. Hedges, & J. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 221–235). New York: Sage.
- Burns, D. J., Hart, J., Kramer, M. E., & Burns, A. D. (2014). Dying to remember, remembering to survive: Mortality salience and survival processing. *Memory*, 22, 36–50. doi: [10.1080/09658211.2013.788660](https://doi.org/10.1080/09658211.2013.788660)
- 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, 823. doi: [10.3389/fpsyg.2014.00823](https://doi.org/10.3389/fpsyg.2014.00823)
- Champely, S. (2009). *pwr: Basic functions for power analysis (R package version 1.1.1)*. Vienna: R Foundation for Statistical Computing.
- Coburn, K. M., & Vevea, J. L. (2016). Weightr: Estimating weight-function models for publication bias in R (R package version 1.0.0). Retrieved from <https://CRAN.R-project.org/package=weightr>
- Cochran, W. G. (1954). The combination of estimates from different experiments. *Biometrics*, 10, 101–129. doi: [10.2307/3001666](https://doi.org/10.2307/3001666)
- Cohen, J. (1965). Some statistical issues in psychological research. In B. B. Wolman (Ed.), *Handbook of clinical psychology* (pp. 95–121). New York: McGraw-Hill.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale: Erlbaum.
- Cooper, H., Hedges, L. V., & Valentine, J. C. (Eds.). (2009). *The handbook of research synthesis and meta-analysis*. New York: Sage.
- Cosmides, L., & Tooby, J. (1994). Beyond intuition and instinct blindness: Toward an evolutionarily rigorous cognitive science. *Cognition*, 50, 41–77. doi: [10.1016/0010-0277\(94\)90020-5](https://doi.org/10.1016/0010-0277(94)90020-5)
- Coursol, A., & Wagner, E. E. (1986). Effect of positive findings on submission and acceptance rates: A note on meta-analysis bias. *Professional Psychology: Research and Practice*, 17, 136–137. doi: [10.1037/0735-7028.17.2.136](https://doi.org/10.1037/0735-7028.17.2.136)
- Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York: Routledge.
- Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, 1, 170–177. doi: [10.1037/1082-989X.1.2.170](https://doi.org/10.1037/1082-989X.1.2.170)
- Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56, 455–463. doi: [10.1111/j.0006-341X.2000.00455.x](https://doi.org/10.1111/j.0006-341X.2000.00455.x)
- Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315, 629–634. doi: [10.1136/bmj.315.7109.629](https://doi.org/10.1136/bmj.315.7109.629)
- Field, A. P., & Gillett, R. (2010). How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology*, 63, 665–694. doi: [10.1348/000711010X502733](https://doi.org/10.1348/000711010X502733)
- Francis, G. (2012). Publication bias and the failure of replication in experimental psychology. *Psychonomic Bulletin & Review*, 19, 975–991. doi: [10.3758/s13423-012-0322-y](https://doi.org/10.3758/s13423-012-0322-y)
- Francis, G. (2014). The frequency of excess success for articles in *Psychological Science*. *Psychonomic Bulletin & Review*, 21, 1180–1187. doi: [10.3758/s13423-014-0601-x](https://doi.org/10.3758/s13423-014-0601-x)
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5, 3–8. doi: [10.3102/0013189X005010003](https://doi.org/10.3102/0013189X005010003)
- Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social science research (Sage Library of Social Research, Vol. 124)*. Beverly Hills: Sage.
- Hedges, L. V. (1982). Estimation of effect size from a series of independent experiments. *Psychological Bulletin*, 92, 490–499. doi: [10.1037/0033-2909.92.2.490](https://doi.org/10.1037/0033-2909.92.2.490)
- Hedges, L., & Olkin, I. (1985). *Statistical models for meta-analysis*. New York: Academic Press.
- Hedges, L. V., & Vevea, J. L. (1998). Fixed-and random-effects models in meta-analysis. *Psychological Methods*, 3, 486–504. doi: [10.1037/1082-989X.3.4.486](https://doi.org/10.1037/1082-989X.3.4.486)
- Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, 327, 557–560. doi: [10.1136/bmj.327.7414.557](https://doi.org/10.1136/bmj.327.7414.557)
- Howe, M. L., & Otgaar, H. (2013). Proximate mechanisms and the development of adaptive memory. *Current Directions in Psychological Science*, 22, 16–22. doi: [10.1177/0963721412469397](https://doi.org/10.1177/0963721412469397)

- Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I² index? *Psychological Methods*, 11, 193–206. doi:10.1037/1082-989X.11.2.193
- Inzlicht, M., Gervais, W., & Berkman, E. (2015). Bias-correction techniques alone cannot determine whether ego depletion is different from zero: Commentary on Carter, Kofler, Forster, & McCullough, 2015. doi:10.2139/ssrn.2659409
- Ioannidis, J. P., & Trikalinos, T. A. (2007). An exploratory test for an excess of significant findings. *Clinical Trials*, 4, 245–253. doi:10.1177/1740774507079441
- John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*, 23, 524–532. doi:10.1177/0956797611430953
- Kelley, K. (2007). Confidence intervals for standardized effect sizes: Theory, application, and implementation. *Journal of Statistical Software*, 20, 1–24. doi:10.18637/jss.v020.i08
- Kruschke, J. K. (2010). What to believe: Bayesian methods for data analysis. *Trends in Cognitive Sciences*, 14, 293–300. doi:10.1016/j.tics.2010.05.001
- Langan, D., Higgins, J. P. T., & Simmonds, M. (2017). Comparative performance of heterogeneity variance estimators in meta-analysis: A review of simulation studies. *Research Synthesis Methods*, 8, 181–198. doi:10.1002/jrsm.1198
- Marín-Martínez, F., & Sánchez-Meca, J. (2010). Weighting by inverse variance or by sample size in random-effects meta-analysis. *Educational and Psychological Measurement*, 70, 56–73. doi:10.1177/0013164409344534
- McBride, D. M., Thomas, B. J., & Zimmerman, C. (2013). A test of the survival processing advantage in implicit and explicit memory tests. *Memory & Cognition*, 41, 862–871. doi:10.3758/s13421-013-0304-y
- Morey, R. D., & Rouder, J. N. (2015). BayesFactor: Computation of Bayes factors for common designs (R package version 0.9.12-2). <https://CRAN.R-project.org/package=BayesFactor>
- Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, 7, 105–125. doi:10.1037/1082-989X.7.1.105
- Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., ... & Contestabile, M. (2015). Promoting an open research culture. *Science*, 348, 1422–1425. doi:10.1126/science.aab2374
- Okada, K. (2013). Is omega squared less biased? A comparison of three major effect size indices in one-way ANOVA. *Behaviormetrika*, 40, 129–147. doi:10.2333/bhmk.40.129
- Okada, K., & Hoshino, T. (2016). Researchers/choice of the number and range of levels in experiments affects the resultant variance-accounted-for effect size. *Psychonomic Bulletin & Review*, 2016. doi:10.3758/s13423-016-1128-0
- Sánchez-Meca, J., & Marín-Martínez, F. (2008). Confidence intervals for the overall effect size in random-effects meta-analysis. *Psychological Methods*, 13, 31–48. doi:10.1037/1082-989X.13.1.31
- Schönbrodt, F. D., Wagenmakers, E.-J., Zehetleitner, M., & Perugini, M. (2017). Sequential hypothesis testing with Bayes factors: Efficiently testing mean differences. *Psychological Methods*, 22, 322–339. doi:10.1037/met0000061
- Schwarzer, G., Carpenter, J. R., & Rücker, G. (2015). An introduction to meta-analysis in R. In *Meta-Analysis With R* (pp. 3–17). Cham, Switzerland: Springer. doi:10.1007/978-3-319-21416-0_1
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22, 1359–1366. doi:10.1177/0956797611417632
- Simonsohn, U. (2017). The funnel plot is invalid because of this crazy assumption: $r(n,d) = 0$. In Data colada. Retrieved April 4, 2017, from <http://datacolada.org/58>
- Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve and effect size: Correcting for publication bias using only significant results. *Perspectives on Psychological Science*, 9, 666–681. doi:10.1177/1745691614553988
- 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, 1146–1152. doi:10.1037/xge0000104
- Smithson, M. J. (2003). *Confidence intervals*. Thousand Oaks: Sage.
- Soderstrom, N. C., & McCabe, D. P. (2011). Are survival processing memory advantages based on ancestral priorities? *Psychonomic Bulletin & Review*, 18, 564–569. doi:10.3758/s13423-011-0060-6
- Stanley, T. D. (2005). Beyond publication bias. *Journal of Economic Surveys*, 19, 309–345. doi:10.1111/j.0950-0804.2005.00250.x
- Stanley, T. D., & Doucouliagos, H. (2013). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods*, 5, 60–78. doi:10.1002/jrsm.1095
- Terrin, N., Schmid, C. H., Lau, J., & Olkin, I. (2003). Adjusting for publication bias in the presence of heterogeneity. *Statistics in Medicine*, 22, 2113–2126. doi:10.1002/sim.1461
- Tooby, J., & Cosmides, L. (2005). Conceptual foundations of evolutionary psychology. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (pp. 5–67). Hoboken: Wiley.
- van Aert, R. C. M. (2017). *Puniform*. GitHub repository, <https://github.com/RobbievanAert/puniform>
- van Aert, R. C. M., 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, 713–729. doi:10.1177/1745691616650874
- van Assen, M. A. L. M., van Aert, R. C. M., & Wicherts, J. M. (2014). Meta-analysis using effect size distributions of only statistically significant studies. *Psychological Methods*, 20, 293–309. doi: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 of Psychology*, 6, 1365. doi:10.3389/fpsyg.2015.01365
- Veroniki, A. A., Jackson, D., Viechtbauer, W., Bender, R., Bowden, J., Knapp, G., ... & Salanti, G. (2016). Methods to estimate the between-study variance and its uncertainty in meta-analysis. *Research Synthesis Methods*, 7, 55–79. doi:10.1002/jrsm.1164
- Vevea, J. L., & Hedges, L. V. (1995). A general linear model for estimating effect size in the presence of publication bias. *Psychometrika*, 60, 419–435. doi:10.1007/bf02294384
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36. doi:10.18637/jss.v036.i03
- Viechtbauer, W., & Cheung, M. W.-L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, 1, 112–125. doi:10.1002/jrsm.11
- Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of *p* values. *Psychonomic Bulletin & Review*, 14, 779–804. doi:10.3758/BF03194105
- Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis*. Beverly Hills: Sage.