Informed Bayesian T-Tests: Online Appendix*

Quentin F. Gronau[1] **, Alexander Ly**[1,2] **and Eric-Jan Wagenmakers** [1]

[1] *University of Amsterdam*
Department of Psychological Methods
Nieuwe Achtergracht 129 B
1018 WT Amsterdam
The Netherlands
e-mail: Quentin.F.Gronau@gmail.com; ej.wagenmakers@gmail.com

[2] *Centrum Wiskunde & Informatica*
Kruislaan 413
PO Box 94079
1090 GB Amsterdam
The Netherlands
e-mail: a.ly@cwi.nl
url: www.alexander-ly.com/; https://jasp-stats.org/

*Correspondence concerning this article should be addressed to: Quentin F. Gronau, University of Amsterdam, Nieuwe Achtergracht 129 B, 1018 WT Amsterdam, The Netherlands. E-mail may be sent to Quentin.F.Gronau@gmail.com. This research was supported by a Netherlands Organisation for Scientific Research (NWO) grant to QFG (406.16.528) and by a Vici grant from the NWO to EJW (016.Vici.170.083), which also funded AL. AL was in part funded by the research program NWO TOP "Safe Bayesian Learning" with project number 617.001.651. Centrum Wiskunde & Informatica (CWI) is the national research institute for mathematics and computer science in the Netherlands. R code can be found on the Open Science Framework: https://osf.io/37vch/.

Appendix A: Technical derivations

Theorem A.1 (The two-sample likelihood based on the grand mean and the effect size parameters). *Let $Y_{ji} \sim \mathcal{N}(\mu_j, \sigma^2)$ for $i = 1, \ldots, n_j$, $j = 1, 2$, $\mu_j = \mu + (-1)^{j+1}\sigma\delta/2$, where μ is the grand mean, δ the standardized effect size, thus, $\varphi = \sigma\delta$ the mean difference, σ the common standard deviation, then the likelihood of the two-sample t-test can be written as*

$$f(d\,|\,\theta) = (2\pi)^{-\frac{n}{2}}\sigma^{-n}\exp\Big\{-\tfrac{n}{2\sigma^2}\big(\mu - [\bar{z} - (q_2 - q_1)]\varphi\big)^2\Big\} \tag{A.1}$$

$$\times \exp\Big\{-\tfrac{1}{2\sigma^2}\big(\nu s_p^2 + n_\delta[\varphi - (\bar{y}_1 - \bar{y}_2)]^2\big)\Big\} \tag{A.2}$$

where $\theta = (\mu, \sigma, \delta)$, and where d refers to the data with a total sample size of $n = n_1 + n_2$, $\nu = n - 2$ degrees of freedom, a pooled sample mean of $\bar{z} = q_1\bar{y}_1 + q_2\bar{y}_2$, where $q_j = n_j/n$, a pooled variance of s_p^2, where $\nu s_p^2 = \sum_{i=1}^{n_1}(y_{1i} - \bar{y}_1)^2 + \sum_{i=1}^{n_2}(y_{2i} - \bar{y}_2)^2$ is the pooled sums of squares, and an effective sample size of $n_\delta = (n_1^{-1} + n_2^{-1})^{-1} = (n_1 n_2)/n$. ◇

Proof. To isolate the dependence of the likelihood on the grand mean μ, we complete the square. To simplify matters we temporarily express the likelihood in terms of the population mean difference $\varphi = \sigma\delta$ and sometimes we use $\xi = \varphi/2$. First recall that the likelihood for each group can be written in terms of its sample mean $\bar{y}_j$ and sums of squares $n_j s_j^2 = \sum_{i=1}^{n_j}(y_{ji} - \bar{y}_j)^2$ for $j = 1, 2$. Hence, since $\nu s_p^2 = n_1 s_1^2 + n_2 s_2^2$, the likelihood of $f(d\,|\,\theta)$ is

$$\begin{aligned} f(d\,|\,\theta) =& (2\pi)^{-\frac{n}{2}}\sigma^{-n}\exp(-\tfrac{n_1}{2\sigma^2}[(\bar{y}_1 - \mu_1)^2 + s_1^2])\exp(-\tfrac{n_2}{2\sigma^2}[(\bar{y}_2 - \mu_2)^2 + s_2^2]) \\ =& (2\pi)^{-\frac{n}{2}}\sigma^{-n}\exp(-\tfrac{1}{2\sigma^2}\nu s_p^2)\exp(-\tfrac{n}{2\sigma^2}\tilde{Q}(d\,|\,\xi, \mu)) \end{aligned} \tag{A.3}$$

where the quadratic function $\tilde{Q}(d\,|\,\xi, \mu)$, which by the simplification $\xi = \varphi/2$ is

$$\begin{aligned} \tilde{Q}(d\,|\,\xi, \mu) =& n_1\big[\mu^2 - 2(\bar{y}_1 - \xi)\mu + \xi^2 - 2\bar{y}_1\xi + \bar{y}_1^2\big] \\ &+ n_2\big[\mu^2 - 2(\bar{y}_2 + \xi)\mu + \xi^2 + 2\bar{y}_2\xi + \bar{y}_2^2\big]. \end{aligned} \tag{A.4}$$

Writing $\bar{z} = q_1\bar{y}_1 + q_2\bar{y}_2$ for the overall mean, the quadratic simplifies to

$$\tilde{Q}(d\,|\,\xi, \mu) = \underbrace{\mu^2 - 2(\bar{z} + [q_2 - q_1]\xi)\mu}_{g(d\,|\,\mu)} + \xi^2 - 2[q_1\bar{y}_1 - q_2\bar{y}_2]\xi + q_1\bar{y}_1^2 + q_2\bar{y}_2^2. \tag{A.5}$$

The function $g(d\,|\,\mu)$ involves the square that needs completing, and with $\eta = \bar{z} + [q_2 - q_1]\xi$ is given by $g(d\,|\,\mu) = (\mu - \eta)^2 - (\bar{z} + [q_2 - q_1]\xi)^2$. Hence,

$$\tilde{Q}(d\,|\,\xi, \mu) = (\mu - \eta)^2 + Q(d\,|\,\xi), \tag{A.6}$$

where

$$Q(d\,|\,\xi) = [1 - (q_2 - q_1)^2]\xi^2 - 2[q_1\bar{y}_1 - q_2\bar{y}_2 + (q_2 - q_1)\bar{z}]\xi + q_1\bar{y}_1^2 + q_2\bar{y}_2^2 - \bar{z},$$

is a quadratic function that does not depend on the grand mean μ. Multiple uses of $\bar{z} = q_1\bar{y}_1 + q_2\bar{y}_2$ and $q_1 + q_2 = 1$ show that $Q(d \mid \xi)$ contains a common factor $q_1 q_2$, that is,

$$Q(d \mid \xi) = 4q_1q_2\xi^2 - 2(2q_1q_2[\bar{y}_1 - \bar{y}_2])\xi + q_1q_2(\bar{y}_1 - \bar{y}_2)^2. \tag{A.7}$$

Using the definition $\xi = \varphi/2$, we see that the constants (4 and 2) in the first two terms of $Q(d \mid \xi)$ cancel and another application of completing the square with respect to φ shows that $Q(d \mid \varphi) = q_1q_2\big(\varphi - [\bar{y}_1 - \bar{y}_2]\big)^2$. Tidying up all terms yields the assertion. □

Corollary A.1.1 (Jeffreys's parameterization-invariant prior on the grand mean)**.** *Let $Y_{ji} \sim \mathcal{N}(\mu_j, \sigma^2)$ for $i = 1, \ldots, n_j$, $j = 1, 2$, where $\mu_j = \mu + (-1)^{j+1}\sigma\delta/2$, then with Jeffreys's parameterization invariant prior $\pi(\mu) \propto 1$ the likelihood reduces to*

$$f(d \mid \sigma, \delta) = (2\pi)^{\frac{1-n}{2}}\sigma^{1-n}n^{-\frac{1}{2}}\exp\Big\{-\tfrac{1}{2\sigma^2}\big(\nu s_p^2 + n_\delta[\varphi - (\bar{y}_1 - \bar{y}_2)]^2\big)\Big\}, \tag{A.8}$$

which is a direct consequence of the normal integral. ◇

If a normal prior on the grand mean μ is used instead, then we would end up with $\bar{z}$ in the reduced likelihood. As the two-sample t-test is concerned with testing whether two population means are the same, we believe that the observed overall sample mean $\bar{z}$ over the two groups is not informative, and should therefore not be part of the two-sample test.

Corollary A.1.2 (Marginal likelihood of the null model)**.** *Let $Y_{ji} \sim \mathcal{N}(\mu_j, \sigma^2)$ for $i = 1, \ldots, n_j$, $j = 1, 2$, where $\mu_j = \mu + (-1)^{j+1}\sigma\delta/2$. Under the null that states that $\delta = 0$, we have $\mu_1 = \mu_2 = \mu$ which we already marginalized out of the likelihood. For the marginal likelihood of the null model, we have to choose a prior on the standard deviation σ for which we use Jeffreys's parameterization invariant prior $\pi(\sigma) \propto \sigma^{-1}$, this leads to*

$$p(d \mid \mathcal{M}_0) = \frac{\Gamma(\frac{\nu+1}{2})}{2\pi^{\frac{\nu+1}{2}}(\nu+2)^{\frac{1}{2}}}\big[\nu s_p^2 + n_\delta(\bar{y}_1 - \bar{y}_2)^2\big]^{-\frac{\nu+1}{2}} \tag{A.9}$$

which follows directly from the gamma integral after a change of variables. ◇

Corollary A.1.3 (Relating the marginal likelihood of the null model to the t-distribution)**.** *Let $Y_{ji} \sim \mathcal{N}(\mu_j, \sigma^2)$ for $i = 1, \ldots, n_j$, $j = 1, 2$, where $\mu_j = \mu + (-1)^{j+1}\sigma\delta/2$, then with the right Haar prior $\pi(\mu, \sigma) \propto \sigma^{-1}$, we get*

$$p(d \mid \mathcal{M}_0) = \begin{cases} \big(2|\bar{y}_1 - \bar{y}_2|\big)^{-1} & d_{\nu<\min} \\ \frac{\Gamma(\frac{\nu+1}{2})}{2\pi^{\frac{\nu+1}{2}}\sqrt{\nu+2}}\big(n_\delta(\bar{y}_1 - \bar{y}_2)^2\big)^{-\frac{\nu+1}{2}} & d_{\text{info},\nu} \\ \frac{\Gamma(\frac{\nu+1}{2})}{2(\pi\nu s_p^2)^{\frac{\nu+1}{2}}\sqrt{\nu+2}}\big(1 + \frac{t^2}{\nu}\big)^{-\frac{\nu+1}{2}} & n_1, n_2, \nu \geq 1 \end{cases} \tag{A.10}$$

where the $d_{\nu<\min}$ *refers to completely uninformative data sets with* $n_1 = n_2 = 1$, $\bar{y}_1 \neq \bar{y}_2$ *and, therefore,* $n_\delta = 1/2$ *and* $\nu s_p^2 = 0$, *while* $d_{\text{info},\nu}$ *refers to an overwhelmingly informative data set with* $\nu s_p^2 = 0$, $\bar{y}_1 \neq \bar{y}_2$ *of sufficient size, as* $n_\delta > 1/2$, $n_1, n_2, \nu \geq 1$. *Note that the typical case with* $n_1, n_2, \nu \geq 1$ *shows much resemblance to its frequentist counterpart, as*

$$p(d \mid \mathcal{M}_0) = C(d) T_\nu(t) \tag{A.11}$$

where $C(d)$ *is a data dependent term that will also appear in the reduced likelihood of* $\mathcal{M}_1$, *that is,*

$$C(d) = \frac{\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})}{2(\pi \nu s_p^2)^{\frac{\nu+1}{2}} \sqrt{\nu + 2}}, \tag{A.12}$$

and $T_\nu(t)$ *denotes the density of a t-distribution centered at zero with scale one and* ν *degrees of freedom.* ◇

When the standardized effect size δ is not necessary zero, the computations are only a bit more complicated. To simplify matters, the reduced likelihood $f(d \mid \delta, \sigma)$ with the grand mean integrated out, Eq. (A.8), is rewritten as

$$f(d \mid \delta, \sigma) = (2\pi)^{\frac{1-n}{2}} n^{-\frac{1}{2}} e^{-\frac{n_\delta}{2}\delta^2} \sigma^{1-n} \tag{A.13}$$

$$\times \exp\Big(- \tfrac{1}{2}\big[n_\delta (\bar{y}_1 - \bar{y}_2)^2 + \nu s_p^2\big]\sigma^{-2} + n_\delta(\bar{y}_1 - \bar{y}_2)\delta\sigma^{-1}\Big). \tag{A.14}$$

For the integral with respect to the Jeffreys's prior on the standard deviation we use a lemma distilled from the Bateman project (Bateman et al., 1954, 1953; Ly et al., 2018).

Theorem A.2 (The reduced likelihood with a right Haar prior on μ, σ)**.** *Let* $Y_{ji} \sim \mathcal{N}(\mu_j, \sigma^2)$, *where* $\mu_j = \mu + (-1)^{j+1}\sigma\delta/2$ *for* $j = 1, 2$, *then with the right Haar prior on the nuisance parameters, that is,* $\pi(\mu, \sigma) \propto \sigma^{-1}$ *the reduced likelihood is*

$$f(d \mid \delta) = 2^{-1} \pi^{\frac{1-n}{2}} n^{-\frac{1}{2}} \Big(n_\delta [\bar{y}_1 - \bar{y}_2]^2 + \nu s_p^2\Big)^{-\frac{\nu+1}{2}} e^{-\frac{n_\delta}{2}\delta^2} \Big[A(d \mid \delta) + B(d \mid \delta)\Big]$$

where

$$A(d \mid \delta) = \Gamma(\tfrac{\nu+1}{2}) {}_1F_1\Big(\tfrac{\nu+1}{2}\,;\, \tfrac{1}{2}\,;\, \tfrac{n_\delta^2 [\bar{y}_1 - \bar{y}_2]^2 \delta^2}{2(\nu s_p^2 + n_\delta [\bar{y}_1 - \bar{y}_2]^2)}\Big) \tag{A.15}$$

$$B(d \mid \delta) = \tfrac{\sqrt{2}(n_\delta [\bar{y}_1 - \bar{y}_2])}{\sqrt{\nu s_p^2 + n_\delta [\bar{y}_1 - \bar{y}_2]^2}} \delta \Gamma(\tfrac{\nu+2}{2}) {}_1F_1\Big(\tfrac{\nu+2}{2}\,;\, \tfrac{3}{2}\,;\, \tfrac{n_\delta^2 [\bar{y}_1 - \bar{y}_2]^2 \delta^2}{2(\nu s_p^2 + n_\delta [\bar{y}_1 - \bar{y}_2]^2)}\Big), \tag{A.16}$$

where ${}_1F_1(a\,;\, b\,;\, z) = \sum_{n=1}^\infty \frac{(a)_n}{(b)_n n!} z^n$ *is the confluent hypergeometric function, where* $(a)_n = \Gamma(a + n)/\Gamma(a)$ *is the Pochhammer symbol for rising factorials. Observe that whenever* $\nu s_p^2 \neq 0$, *we can divide the arguments of the confluent hypergeometric function and the second term that requires to be taken to power of* $-\frac{\nu+1}{2}$ *which leads to*

$$f(d \mid \delta) = C(d) T_\nu(t \mid \sqrt{n_\delta}\delta) \tag{A.17}$$

where $C(d)$ is exactly the same dependent term that appeared in $p(d \,|\, \mathcal{M}_0)$, and $T_\nu(t \mid a)$ refers to the density of a t-distribution at t with ν degrees of freedom and non-centrality parameter a. ⋄

Proof. This follows directly from the lemma distilled from the Bateman project (Ly et al., 2018). □

Corollary A.2.1 (Predictive matching)**.** *Let $Y_{ji} \sim \mathcal{N}(\mu_j, \sigma^2)$ for $i = 1, \ldots, n_j$ for $j = 1, 2$ and let $d_{\nu < \min}$ refer to a completely uninformative data set with $n_1 = n_2 = 1$ and $\bar{y}_1 \neq \bar{y}_2$, thus, $\nu = 0$, $n_\delta = 1/2$ and automatically $\nu s_p^2 = 0$. With a right Haar prior on the nuisance parameters, that is, $\pi(\mu, \sigma) \propto \sigma^{-1}$ the reduced likelihood is then*

$$f(d_{\nu<\min} \,|\, \delta) = (2|\bar{y}_1 - \bar{y}_2|)^{-1}\big[1 + \text{sign}(\bar{y}_1 - \bar{y}_2)\text{Erf}(\tfrac{\delta}{2})\big], \tag{A.18}$$

where $\text{Erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} \text{d}t$ refers to the error function. As such, for a predictively matched Bayes factor we require that the prior $\pi(\delta)$ is proper and symmetric around zero. ⋄

Proof. Crucial to the understanding of the result is the realization that the error function is an odd function of δ and note that for the uninformative data set, the reduced likelihood

$$f(d_{n_\delta \leq .5} \,|\, \delta) = p(d_{n_\delta \leq .5} \,|\, \mathcal{M}_0) + p(d_{n_\delta \leq .5} \,|\, \mathcal{M}_0)\text{sign}(\bar{y}_1 - \bar{y}_2)\text{Erf}(\tfrac{\delta}{2}). \tag{A.19}$$

As such,

$$p(d_{n_\delta \leq .5} \,|\, \mathcal{M}_1) = p(d_{n_\delta \leq .5} \,|\, \mathcal{M}_0) + \text{sign}(\bar{y}_1 - \bar{y}_2) \int \text{Erf}(\tfrac{\delta}{2})\pi(\delta)\text{d}\delta \tag{A.20}$$

and the latter term is zero when $\pi(\delta)$ is even. In other words, for the uninformative data set we have $p(d_{n_\delta \leq .5} \,|\, \mathcal{M}_0) = p(d_{n_\delta \leq .5} \,|\, \mathcal{M}_1)$ and therefore $\text{BF}_{10}(d_{n_\delta \leq .5}) = 1$. □

Corollary A.2.2 (Behaviour of the reduced likelihood for informative data)**.** *Let $Y_{ji} \sim \mathcal{N}(\mu_j, \sigma^2)$ for $i = 1, \ldots, n_j$ for $j = 1, 2$. For informative data $d_{\text{info},\nu}$ with $\nu \geq 1$, $n_\delta \geq 1/2$ (meaning there is at least one observation per group), $\nu s_p^2 = 0$ and $\bar{y}_1 \neq \bar{y}_2$, then*

$$f(d_{\text{info},\nu} \,|\, \delta) \sim \delta^\nu \tag{A.21}$$

as $\delta \to \infty$. In other words, for δ large, the reduced likelihood behaves like a polynomial in δ of order ν. ⋄

Proof. Using the fact that for $x \to \infty$

$${}_1F_1(a\,;\,b\,;\,x) \sim e^x x^{a-b}/\Gamma(a) \tag{A.22}$$

we see that

$$e^{-\frac{n_\delta \delta^2}{2}} A(d_{\text{info},\nu} \,|\, \delta) \sim 2^{-\frac{\nu}{2}} (\sqrt{n_\delta}\delta)^\nu \tag{A.23}$$

$$e^{-\frac{n_\delta \delta^2}{2}} B(d_{\text{info},\nu} \,|\, \delta) \sim 2^{-\frac{\nu-1}{2}} \sqrt{2}\text{sign}(\bar{y}_1 - \bar{y}_2)(\sqrt{n_\delta}\delta)^\nu \tag{A.24}$$

where the exponential stems from the first terms of $f(d_{\text{info},\nu} \mid \delta)$. Tidying up all terms now yields the result. □

Corollary A.2.3 (Marginal posterior for δ)**.** *The previous theorems and corollaries imply that whenever $\nu s_p^2 \neq 0$, that*

$$\pi(\delta \mid d) \propto T_\nu(t \mid \sqrt{n_\delta}\delta)\pi(\delta). \tag{A.25}$$

In particular, for the t-prior we have

$$\pi_{\kappa,\mu_\delta,\gamma}(\delta \mid d) \propto T_\nu(t \mid \sqrt{n_\delta}\delta)\tfrac{1}{\gamma}T_\kappa(\tfrac{\delta-\mu_\delta}{\gamma}) \tag{A.26}$$

with normalizing constant $\int T_\nu(t \mid \sqrt{n_\delta}\delta)\tfrac{1}{\gamma}T_\kappa(\tfrac{\delta-\mu_\delta}{\gamma})\mathrm{d}\delta$ that can be easily and accurately computed numerically. ◇

Theorem A.3 (Scaled mixtures of normal priors on the effect size)**.** *Let $Y_{ji} \sim \mathcal{N}(\mu_j, \sigma^2)$, where $\mu_j = \mu + (-1)^{j+1}\sigma\delta/2$ for $j = 1, 2$, then with the right Haar prior on the nuisance parameters, that is, $\pi(\mu, \sigma) \propto \sigma^{-1}$ and a normal prior on δ, thus, $\delta \sim \mathcal{N}(\mu_\delta, g)$ as in Gönen et al. (2005), then whenever $\nu s_p^2 \neq 0$*

$$p_{\mu_\delta,g}(d \mid \mathcal{M}_1) = C(d)\frac{1}{\sqrt{1+n_\delta g}}T_\nu\Big(\frac{t}{\sqrt{1+n_\delta g}}\,;\,\sqrt{\frac{n_\delta}{1+n_\delta g}}\mu_\delta\Big), \tag{A.27}$$

where $C(d)$ is exactly the same dependent term that appeared in $p(d \mid \mathcal{M}_0)$.

We use the notation $p_{\mu_\delta,g}(d \mid \mathcal{M}_1)$ to convey that μ_δ and g are fixed and chosen. On the other hand, when a scaled mixture of normal priors is used on δ, say, $\delta \sim \mathcal{N}(\mu_\delta, g)$, then we still have one integral to compute for the marginal likelihood of $\mathcal{M}_1$ and we then call $f_{\mu_\delta}(d \mid g) = p_{\mu_\delta,g}(d \mid \mathcal{M}_1)$ the reduced likelihood. The marginal likelihood is then

$$p_{\mu_\delta}(d \mid \mathcal{M}_1) = \int f_{\mu_\delta}(d \mid g)\pi(g)\mathrm{d}g. \tag{A.28}$$

Examples of specific $\pi(g)$ in the context of Bayes factors are given in Bayarri et al. (2012), Liang et al. (2008) and Maruyama and George (2011). The latter is a beta prime distribution adopted by Wang and Liu (2016) within the context of the two-sample t-test. ◇

Proof. The trick is once again to complete the squares, followed by tedious computations as were done above. □

Appendix B: Prior Elicitation

Here we describe in more detail how we elicited the prior distribution from Dr. Suzanne Oosterwijk. We applied the "Roulette" method which presents the expert with a grid of equally sized bins that cover the plausible range of possible effect size values. The expert then has to allocate chips to the bins to construct the prior distribution. The MATCH Uncertainty Elicitation Tool provides an immediate feedback by showing a fitted parametric distribution that matches

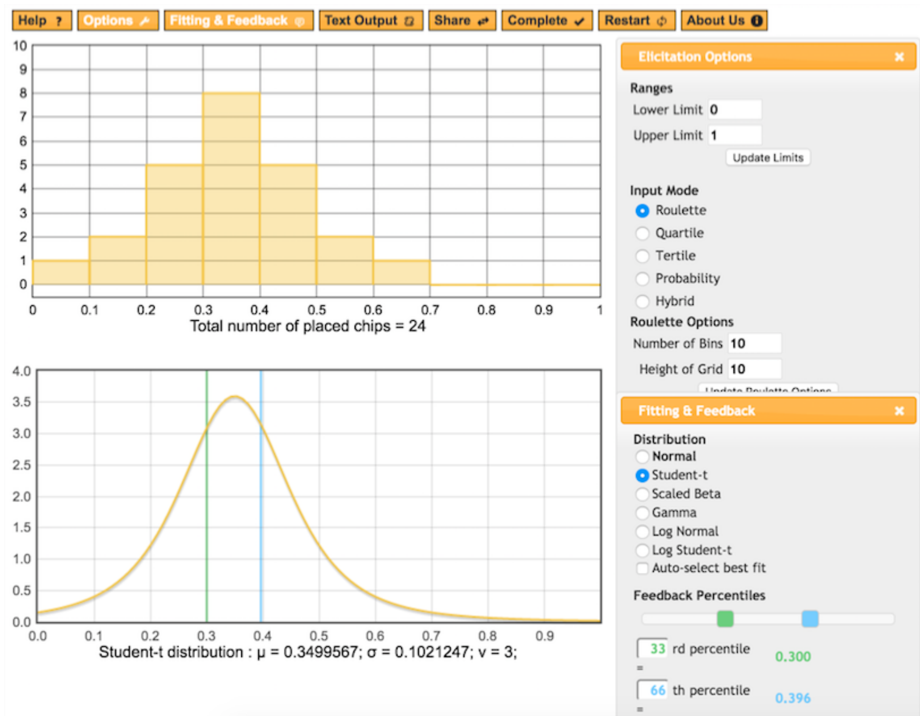

FIG 1. *Depiction of the results of the prior elicitation using the MATCH Uncertainty Elicitation Tool (*`http://optics.eee.nottingham.ac.uk/match/uncertainty.php`*). The top part of the figure displays the results of the "Roulette" method where the expert has to allocate chips to bins in order to elicit an informed prior distribution, the bottom-part shows the fitted t-distribution. The bottom-right corner displays the feedback that the expert received about the implied 33%-tile and 66%-tile values. Figure available at* `https://tinyurl.com/jwqzt7c` *under CC license* `https://creativecommons.org/licenses/by/2.0/`*.*

the distribution specified via the chips (in our case, a t-distribution fit), and it also provides the user with the 33%-tile and 66%-tile values that correspond to the specified distribution. Figure 1 displays a screenshot of the Roulette method with feedback using the MATCH Uncertainty Elicitation Tool. The fitted t-distribution that is shown in Figure 1, that is, a t-distribution with location 0.350, scale 0.102, and degrees of freedom 3 is also the final result of the prior elicitation. Since the MATCH Uncertainty Elicitation Tool only allows fitting t-distributions with three degrees of freedom, we also presented the expert with t-distributions with different number of degrees of freedom using R (R Core Team, 2016). However, the expert decided that the t-distribution with three degrees of freedom captured her prior knowledge best.

Appendix C: Reanalysis of all 17 facial feedback replication studies

Fig. 2 displays the (nondirectional) posterior distributions based on the informed $\frac{1}{0.102}T_3\left(\frac{\delta-0.350}{0.102}\right)$ prior and the default Cauchy$(0, 1/\sqrt{2})$ prior for the reanalysis of all 17 replication studies. The posterior distributions for δ differ noticeably for the informed and the default prior specification. Based on the informed prior, the posterior distributions are shifted towards larger effect size values and the posteriors are more peaked than the ones based on the default prior.

Fig. 3 displays the (directional) informed and default Bayes factors for the reanalysis of all 17 replication studies. Each dot corresponds to a study; the x-coordinate corresponds to the one-sided informed Bayes factor and the y-coordinate corresponds to the one-sided default Bayes factor as reported in Wagenmakers et al. (2016). The results are qualitatively similar. For Bayes factors smaller than about six, the one-sided default Bayes factor provides slightly more evidence for the null hypothesis than the one-sided informed Bayes factor. For Bayes factors larger than about six, this pattern is reversed.

In sum, for parameter estimation it matters whether the analysis is based on an informed prior distribution or a default prior distribution. In contrast, for hypothesis testing the results are relatively similar: both the informed Bayes factor and the default Bayes factor support the conclusion that the original study by Strack et al. (1988) could not be successfully replicated by the 17 labs involved in the replication attempt.

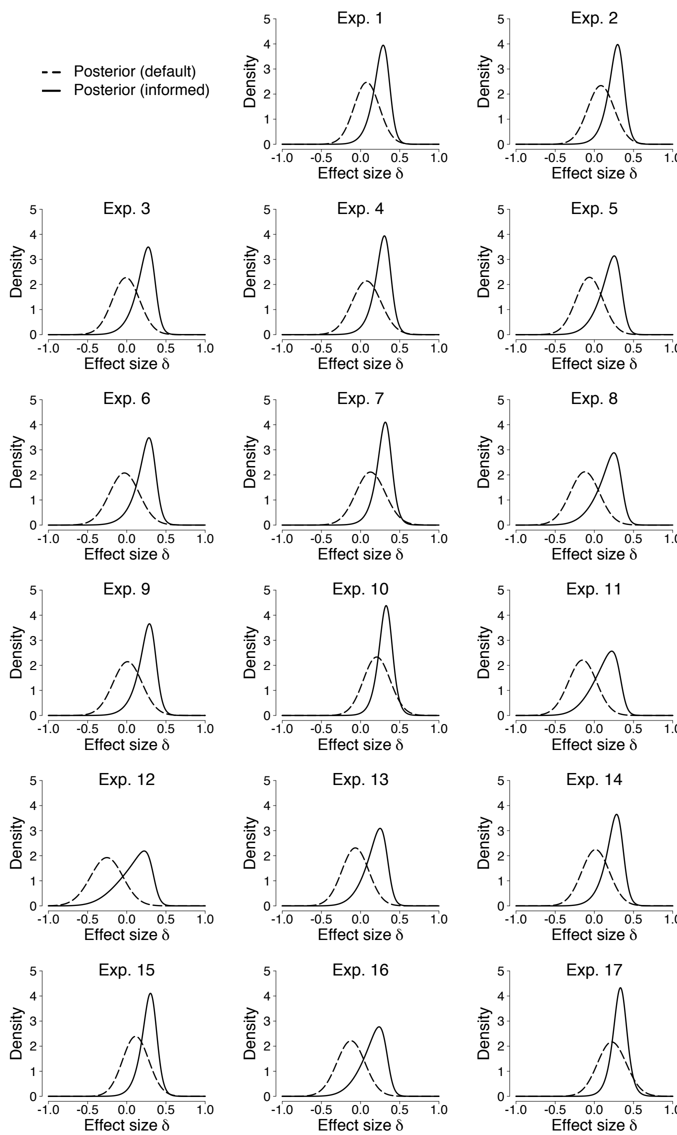

FIG 2. *Comparison of the posterior distributions for δ based on the informed t-distribution with location 0.350, scale 0.102, and three degrees of freedom and based on the zero-centered default Cauchy prior with scale $1/\sqrt{2}$ for the facial feedback hypothesis replication data from the 17 labs. Figure available at* https://tinyurl.com/l7pxbno *under CC license* https://creativecommons.org/licenses/by/2.0/.

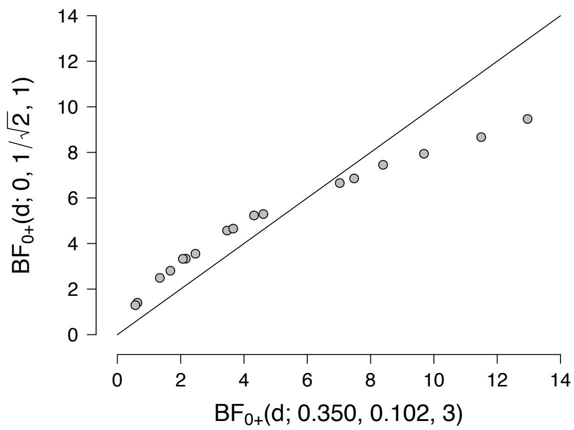

FIG 3. *Comparison of the one-sided default and one-sided informed Bayes factor analysis of the facial feedback hypothesis replication data from the 17 labs. The x-coordinate corresponds to the one-sided informed Bayes factor and the y-coordinate corresponds to the one-sided default Bayes factor as reported in Wagenmakers et al. (2016). Figure available at https://tinyurl.com/ke4489k under CC license https://creativecommons.org/licenses/by/2.0/.*

References

Bateman, H., Erdélyi, A., Magnus, W., Oberhettinger, F., Tricomi, F. G., Bertin, D., Fulks, W. B., Harvey, A. R., Thomsen Jr, D. L., Weber, M. A., Whitney, E. L., and Stampfel, R. (1953). *Higher transcendental functions*, volume 2. McGraw-Hill New York.

Bateman, H., Erdélyi, A., Magnus, W., Oberhettinger, F., Tricomi, F. G., Bertin, D., Fulks, W. B., Harvey, A. R., Thomsen Jr, D. L., Weber, M. A., Whitney, E. L., and Stampfel, R. (1954). *Tables of integral transforms*, volume 1. McGraw-Hill.

Bayarri, M. J., Berger, J. O., Forte, A., and García-Donato, G. (2012). Criteria for Bayesian model choice with application to variable selection. *The Annals of Statistics*, 40(3):1550–1577.

Gönen, M., Johnson, W. O., Lu, Y., and Westfall, P. H. (2005). The Bayesian two-sample t test. *The American Statistician*, 59(3):252–257.

Liang, F., Paulo, R., Molina, G., Clyde, M. A., and Berger, J. O. (2008). Mixtures of g priors for Bayesian variable selection. *Journal of the American Statistical Association*, 103(481):410–423.

Ly, A., Marsman, M., and Wagenmakers, E.-J. (2018). Analytic posteriors for Pearson's correlation coefficient. *Statistica Neerlandica*, 72(1):4–13.

Maruyama, Y. and George, E. I. (2011). Fully Bayes factors with a generalized g-prior. *The Annals of Statistics*, 39(5):2740–2765.

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Strack, F., Martin, L. L., and Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: A nonobtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology*, 54:768–777.

Wagenmakers, E.-J., Beek, T., Dijkhoff, L., Gronau, Q. F., Acosta, A., Adams, R., Albohn, D. N., Allard, E. S., Benning, S. D., Blouin–Hudon, E.-M., Bulnes, L. C., Caldwell, T. L., Calin–Jageman, R. J., Capaldi, C. A., Carfagno, N. S., Chasten, K. T., Cleeremans, A., Connell, L., DeCicco, J. M., Dijkstra, K., Fischer, A. H., Foroni, F., Hess, U., Holmes, K. J., Jones, J. L. H., Klein, O., Koch, C., Korb, S., Lewinski, P., Liao, J. D., Lund, S., Lupiáñez, J., Lynott, D., Nance, C. N., Oosterwijk, S., Özdoğru, A. A., Pacheco–Unguetti, A. P., Pearson, B., Powis, C., Riding, S., Roberts, T.-A., Rumiati, R. I., Senden, M., Shea–Shumsky, N. B., Sobocko, K., Soto, J. A., Steiner, T. G., Talarico, J. M., van Allen, Z. M., Vandekerckhove, M., Wainwright, B., Wayand, J. F., Zeelenberg, R., Zetzer, E. E., and Zwaan, R. A. (2016). Registered Replication Report: Strack, Martin, & Stepper (1988). *Perspectives on Psychological Science*, 11:917–928.

Wang, M. and Liu, G. (2016). A simple two-sample Bayesian t-test for hypothesis testing. *The American Statistician*, 70(2):195–201.