¹ Re-analysing the data from Moffatt et al. (2020): A textbook illustration of the absence of

² evidence fallacy

³

⁴ Submitted to Meta-Psychology. Click here to follow the fully transparent editorial process

⁵ of this submission. Participate in open peer review by commenting through hypothes.is

⁶ directly on this preprint.

⁷ Ladislas Nalborczyk[1]

⁸ [1] Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, 38000 Grenoble, France

⁹ Author Note

¹⁰ Correspondence concerning this article should be addressed to Ladislas Nalborczyk,

¹¹ GIPSA-lab, CNRS, Univ. Grenoble Alpes, 11 Rue des Mathématiques, 38400

¹² Saint-Martin-d'Hères, France. E-mail: ladislas.nalborczyk@gipsa-lab.fr

Abstract

Moffatt et al. (2020) reported the results of an experiment (N = 26 in the final sample) comparing the facial (surface) electromyographic correlates of mental rumination and distraction, following an experimentally induced stressor. Based on the absence of significant difference in the perioral muscular activity between the rumination and distraction conditions, Moffatt et al. (2020) concluded that *self-reported* inner experience was unrelated to peripheral muscular activity as assessed using surface electromyography. We suggest this conclusion is hasty and based on waggly evidence. Indeed, concluding on the absence of an effect based on a low-powered non-significant p-value is strongly problematic/uninformative. Moreover, the relation between self-reports and physiological measures was not *directly* assessed, but only indirectly inferred from differences (or absence thereof) in group means. Given the ample inter-individual variability in these measures (as suggested by our reanalysis), we think inferring the individual-level relation between self-reports and physiological measures from group means is inappropriate. Given these limitations, we conclude that it is unclear whether the target article adds to the current/extent knowledge and we suggest ways forward, both from a theoretical and from a methodological perspective. Complete source code, reproducible analyses, and figures are available at https://github.com/lnalborczyk/inner_experience_EMG.

Keywords: NHST, Bayesian, fallacy, reanalysis, inner speech, rumination, electromyography

Introduction

The activity of silently talking to oneself or "inner speech" is a foundational ability, allowing oneself to remember, plan self-motivate or self-regulate. Despite its multiple adaptive functions in everyday life, inner speech can go awry and leads to sustained negative… These inner speech "dysfunctions" (for reviews, see Alderson-Day & Fernyhough, 2015; Lœvenbruck et al., 2018; Perrone-Bertolotti et al., 2014)…

Given the predominantly verbal nature of rumination (e.g., Ehring & Watkins, 2008; Goldwin et al., 2013; Goldwin & Behar, 2012; McLaughlin et al., 2007), we previously proposed to consider rumination as a form of inner speech and to study it using the methods that have been used historically to study other forms of inner speech, namely, by using surface electromyography and motor interference protocols (e.g., Nalborczyk et al., 2017; Nalborczyk, 2019; Nalborczyk, Perrone-Bertolotti, et al., 2020; Nalborczyk, Banjac, et al., 2020). We first showed that induced rumination was accompanied by increased facial (both over a forehead and a perioral site) muscular activity in comparison to a rest period (Nalborczyk et al., 2017). However, because the rumination condition did not have a proper control condition in this first study, it was unclear whether this perioral activity were specifically related to (inner) speech processes. Therefore, we ran an extension of this study, in which we compared verbal to non-verbal rumination, which suggested that the facial EMG correlates we have previously identified were not specifically related to the verbal content of the ruminative thoughts (Nalborczyk, Banjac, et al., 2020). We discussed these findings in length and proposed several theoretical interpretations that can account for these results in the discussion section of Nalborczyk, Banjac, et al. (2020) and more extensively in Nalborczyk (2019). Although these discussions were blatantly ignored by Moffatt et al. (2020), their experimental design nevertheless had the potential to inform our understanding of the involvement of the speech motor system in different varieties of inner speech as well as to clarify the relation between the peripheral correlates of inner

⁵⁹ speech and the (self-reported) subjective experience.

⁶⁰   The main conclusion from Moffatt et al. (2020) is that inner experience between

⁶¹ induced rumination and distraction differs "without a change in electromyographic

⁶² correlates of inner speech". In other words, they suggest that the subjective experience of

⁶³ inner speech is unrelated (or loosely related) to the electromyographic correlates of inner

⁶⁴ speech, which are thought to be represented mostly by the EMG amplitude recorded over

⁶⁵ the OOI and OOS muscles. However, for this in-sample observation to be of interest in an

⁶⁶ out-of-sample context (i.e., to be informative of other non-observed individuals, or said

⁶⁷ otherwise, to bring information about the population), this absence of difference has to be

⁶⁸ based on sufficiently powered sample size (given the target effect size) as well as on reliable

⁶⁹ measures. This is unlikely to be the case here, for reasons that we will present and discuss

⁷⁰ in the following. Moreover, a simple visual exploration of the data reveals important

⁷¹ variability between individuals in the main effect of interest. That is, some participants

⁷² had higher perioral (OOS and OOI) muscular activity in the rumination condition than in

⁷³ the distraction condition, and some other participants showed the reverse pattern. This

⁷⁴ suggests unexplored variation in the determinants of this effect (e.g., the content of the

⁷⁵ inner experience). Indeed, the relation between the inner experience and the physiological

⁷⁶ correlates of inner speech production was only inferred from group means. However, given

⁷⁷ the previous point, this appears highly problematic. We explore each of these limitations

⁷⁸ and suggests ways forward in the following section.

⁷⁹   **Exploring the data**

⁸⁰   Moffatt et al. (2020) recorded... in 26 participants (data available at

⁸¹ https://osf.io/hj7tz/)... The EMG data is depicted in Figure 1 by condition (where `BAS`,

⁸² `DIS`, and `RUM` refer to the baseline, distraction, and rumination conditions, respectively)

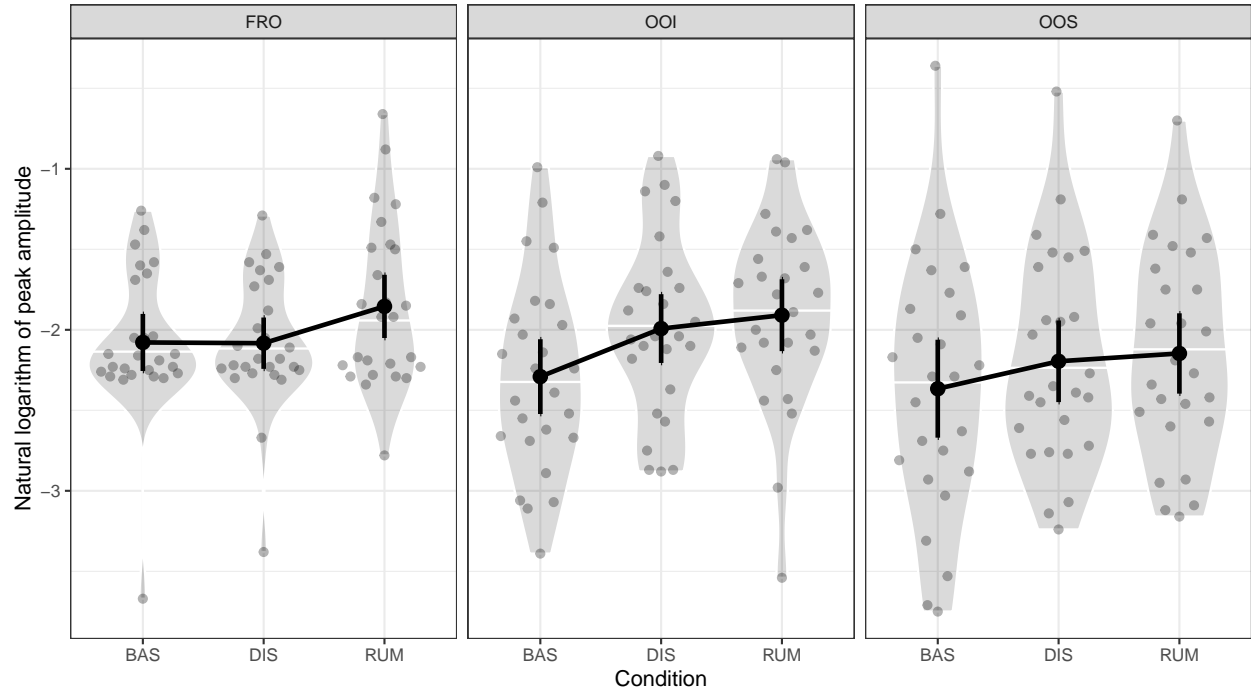⁸³ and by muscle (`FRO`, `OOI`, `OOS`).

Figure 1. Average natural logarithm of the EMG peak amplitude per muscle and condition. The black dots and intervals represent the by-group average and 95% confidence interval (N = 26). The horizontal white line in the violin plot represents the median. The grey dots represent the individual-level average natural logarithm of the EMG amplitude by muscle and condition.

84   We fitted a multivariate Bayesian regression model with varying-intercepts (by

85 participant) and weakly informative priors on these data using the `brms` package (Bürkner,

86 2017)...

87   Posterior distribution of the difference between the distraction and rumination

88 conditions...

89   A summary of the estimations from this model is presented in Table 1. For each

90 muscle, the intercept gives the estimated of the natural logarithm of the EMG peak

91 amplitude in the baseline condition, whereas the `conditionDIS` and `conditionRUM`

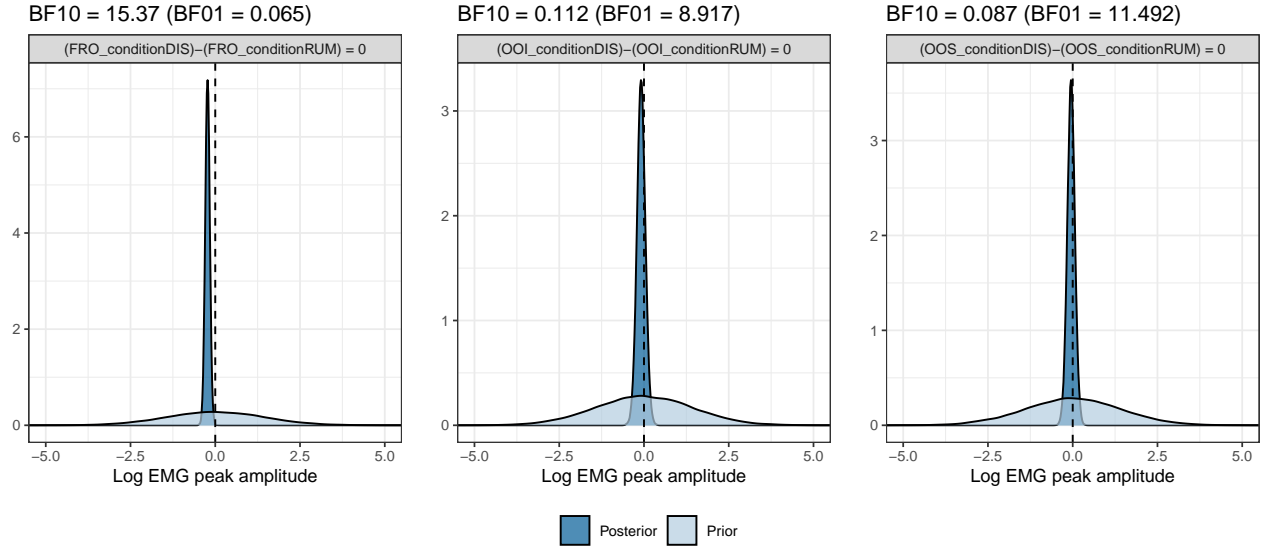92 parameters indicate deviations from the baseline in the distraction and rumination

Figure 2. Savage-Dickey Bayes factor for the difference between the rumination and distraction conditions for each muscle. The BF is computed as the ratio of the posterior density to the prior density at $\theta = 0$.

conditions (respectively).

Table 1

Estimated value of the natural logarithm of the EMG peak amplitude in each condition and for each muscle.

Term	Estimate	SE	Lower	Upper	Rhat	BF10
FRO_Intercept	-2.08	0.10	-2.27	-1.89	1.00	NA
OOS_Intercept	-2.36	0.14	-2.64	-2.09	1.00	NA
OOI_Intercept	-2.29	0.12	-2.52	-2.05	1.00	NA
FRO_conditionDIS	0.00	0.07	-0.14	0.13	1.00	0.06
FRO_conditionRUM	0.22	0.07	0.09	0.36	1.00	10.51
OOS_conditionDIS	0.16	0.11	-0.06	0.38	1.00	0.35
OOS_conditionRUM	0.21	0.11	-0.01	0.43	1.00	0.71
OOI_conditionDIS	0.29	0.12	0.06	0.53	1.00	2.31
OOI_conditionRUM	0.37	0.12	0.14	0.61	1.00	15.22

Note. For each effect, the 'Estimate' reports the estimated average value of the natural logarithm of the EMG peak amplitude, followed by its standard error (SE). The 'Lower' and 'Upper' columns contain the lower and upper bounds of the 95% CrI, whereas the 'Rhat' column reports the Gelman-Rubin statistic. The last column reports the BF in favour of the alternative hypothesis (relative to the null hypothesis).

94 ...

Concluding on the null from low-powered studies: what could go wrong?

96 There is an infamous tradition of running uninformative null-hypothesis significance
97 tests in Psychology (e.g., Meehl, 1997, 1978, 1990a, 1990b, 1967). By "uninformative", we

98 mean that some null-hypothesis significance tests are often *not* diagnostic with regards to

99 the substantive question of interest…

100   As highlighted by many authors (e.g., Pollard & Richardson, 1987; Rouder et al.,

101 2016), concluding on an absence of difference based on not obtaining evidence for the

102 difference is the continuous extension of the logical fallacy of the… The argument from

103 ignorance, such as "Science has found no proof of intelligent life nearby us in space,

104 therefore intelligent life does not exist nearby us in space."… the absence of evidence fallacy

105 or fallacy of acceptance…

106   This problem is tackled in modern usages of null-hypothesis significance test by

107 ensuring that the test has good *severity* (e.g., Mayo & Spanos, 2006; Mayo, 2018). In

108 general terms, we have evidence for a claim to the extent that it survives a stringent

109 scrutiny, that is if it survives *severe tests*. In other words, some claim (e.g., $\theta = 0$) is said

110 to be *severely tested*) if it had great chances of being falsified, was the claim false. More

111 formally, we can define SEV$(T, x0, H)$, the severity with which claim H passes test T with

112 outcome $x0$, and SEV$(\mu > \mu_1) = \Pr(d(X) \leq d(x0); \mu = \mu_1)$ (Mayo, 2018; Mayo & Spanos,

113 2006)…To put it simply… https://www.analytics-toolkit.com/glossary/severity/…

114   Anticipating the critics on the power of their study (a critic that was probably raised

115 during peer review), Moffatt et al. (2020) report the results of a (possibly ran a posteriori)

116 power analysis using the effect size reported in Nalborczyk et al. (2017) of $d = 0.72$, which

117 is highly optimistic estimate of the substantive effect of interest in the target article (i.e.,

118 the difference in EMG amplitude between the rumination and distraction conditions) as

119 this effects represents the standardised mean difference *between a rest period and a*

120 *rumination one* (Nalborczyk et al., 2017)…

```r
# How many participants do we need for a target statistical power of 0.8?
library(pwr)
```

```
pwr.t.test(

  d = 0.72, sig.level = 0.05, power = 0.8,

  type = "one.sample", alternative = "two.sided"

  )
```

```
## 

##        One-sample t test power calculation

## 

##                 n = 17.16004

##                 d = 0.72

##         sig.level = 0.05

##             power = 0.8

##       alternative = two.sided
```

We suggest the (a priori) power of the study ran by Moffatt et al. (2020) was much lower than suggested by the authors. Indeed, we may speculate that the effect (i.e., the standardised mean difference in EMG amplitude) between the rumination and distraction conditions may be much weaker than the effect (i.e., the standardised mean difference in EMG amplitude) between the rumination and the rest conditions. If we assume that the former is half the size of the latter (which seems reasonable given the high inter-individual variability in such effects, cf. the next section but also Nalborczyk, Grandchamp, et al., 2020), therefore the a priori power of the main statistical test from Moffatt et al. (2020) was around 0.44, meaning that they had less than 1 chance out of 2 to find a significant effect, given that the effect in the population was actually 0.36. Because this is less than the chance of obtaining a head in a coin flip, we feel these resources may have been better invested.

```r
# A priori power for n = 26 (per condition) and d = 0.36
pwr.t.test(
  n = 26, d = 0.72 / 2, sig.level = 0.05,
  type = "one.sample", alternative = "two.sided"
  )
```

```
##
##      One-sample t test power calculation
##
##              n = 26
##              d = 0.36
##      sig.level = 0.05
##          power = 0.4228455
##    alternative = two.sided
```

Anticipating the legitimate critique that the absence of a significant difference is not *necessarily* "significant" evidence for the absence of an effect, Moffatt et al. (2020) reported the following Bayes factor (BF) analysis:

"[…] therefore it is possible that the sample size of the present study lacked sufficient power to detect the effect of rumination on muscle activity. In order to test this, a Bayesian paired samples t-test was conducted for the peak log values of muscle activity between the rumination and distraction conditions. This revealed strong evidence in favour of the alternative hypothesis for the FRO muscle ($B_{10} = 18.79$), and moderate evidence in favour of the null hypothesis for the OOS ($B_{10} = 0.232$) and OOI ($B_{10} = 0.278$) muscles, according to current guidelines for interpreting Bayes factors [43]."

While we appreciate the effort, the current approach poses new problems. First,

161 contrary to what the authors suggest, computing a BF (i.e., comparing two models) does

162 not solve *at all* the problem of low power…

163   We first fitted a multivariate Bayesian regression model with varying-intercepts (by

164 participant) and weakly informative priors on these data using the `brms` package (Bürkner,

165 2017). From there, we i) generated new datasets from the posterior predictive distribution

166 and ii) we computed the BF in favour of the alternative hypothesis (BF_{10}) using the

167 `BayesFactor` package (Morey & Rouder, 2018). We used a "medium" prior (i.e., a scale of

168 1) on the scale of the Cauchy prior for the alternative hypothesis. We repeated this

169 procedure for varying sample sizes from 20 to 200 participants (by increments of 10

170 participants) with 1000 simulations (i.e., 1000 simulated datasets) for each sample size.

171   As shown in Figure 3, the natural logarithm of the BF in favour of the alternative

172 hypothesis is growing proportionally with the sample size. More precisely, whereas low

173 sample sizes (i.e., below 80) support the null hypothesis, adequately-powered sample sizes

174 support the alternative hypothesis for all three facial muscles for sample sizes larger than 80

175 participants. For instance, the average BF_{10} computed for the OOI muscle with a sample

176 size of 160 participants is of $\exp(2.18) \approx 8.85$, indicating that these data are approximately

177 8.85 times more likely under the alternative hypothesis than under the null hypothesis.[1]

178   We should keep in mind some limitations of this analysis, which uses simulated

179 datasets form the posterior predictive distribution estimated on the data collected by

180 Moffatt et al. (2020). This analysis is the loose Bayesian analogue of the frequentist

181 post-hoc power analysis, which has been much criticised (e.g., Lakens, 2014). Most

182 importantly, an assumption of the present analysis is that the data from Moffatt et al.

183 (2020) is our best source of information regarding the main effect of interest (in addition to

184 the prior we specified earlier). However, the present analysis also differs from the

---

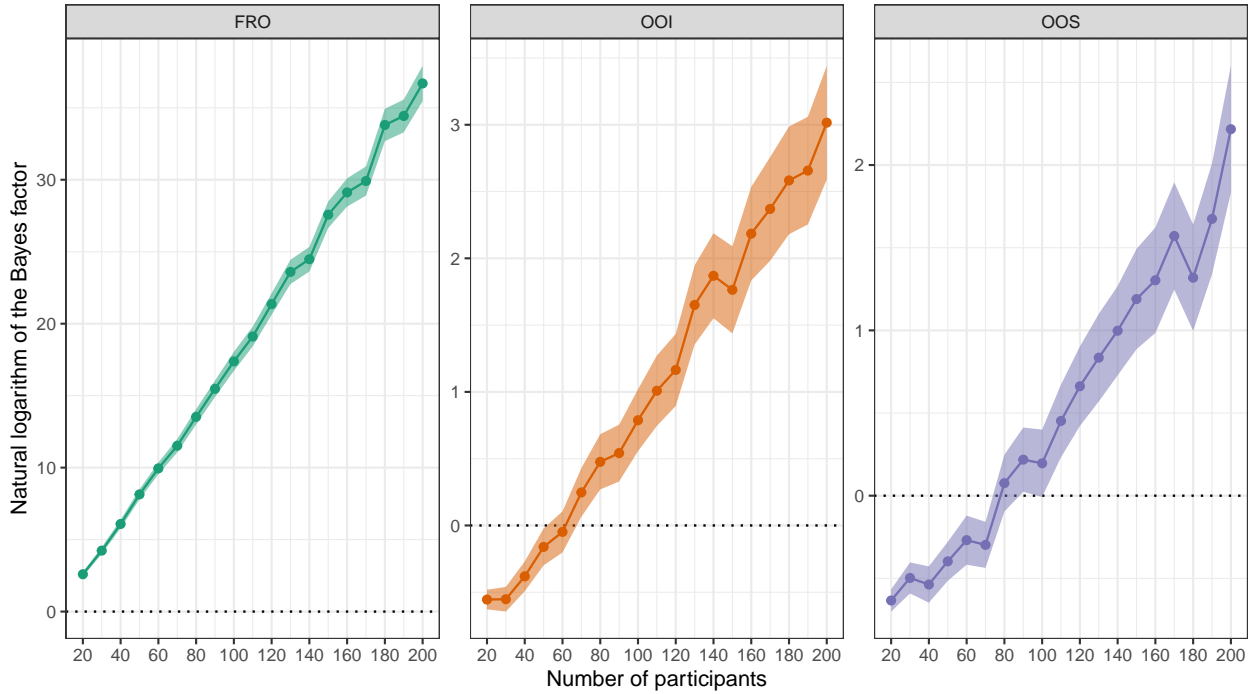[1] Alternatively, the BF can be interpreted as an updating factor, from prior odds to posterior odds.

Figure 3. Average natural logarithm of the Bayes factor in favour of the alternative hypothesis, along with its standard error, computed over 1000 datasets of increasing size simulated from the posterior predictive distribution of the varying-intercept multivariate Bayesian regresion model. A log-BF belows 0 represents evidence for the null hypothesis (relative to the alternative) and a log-BF above 0 represents evidence for the alternative hypothesis (relative to the null).

frequentist post-hoc power analysis on several grounds. First, with the present analysis, we
do not aim to assess the ability of our statistical test to pass some dichotomic threshold
(e.g., accept/reject). Instead, we aim to assess how the BF_{10} (i.e., the evidence for the
alternative hypothesis, relative to the null hypothesis) behaves with varying sample sizes.
Second, the present analysis relies on the posterior predictive distribution of the model
fitted on the data from Moffatt et al. (2020), which naturally incorporates uncertainty
about the effect of interest. By simulating datasets of varying sample sizes from the
posterior predictive distribution (and by relying on a large number of simulations),
uncertainty about the effect size is naturally incorporated into the simulation.

Within-subject manipulation of rumination and distraction

In Nalborczyk, Banjac, et al. (2020), we manipulated the modality of rumination (whether it is verbal or non-verbal) in a between-subject manner to avoid order effects... In contrast to this approach, Moffatt et al. (2020) asked participants to ruminate and then distract themselves (or reciprocally), after an induced stressor (an induced failure)...
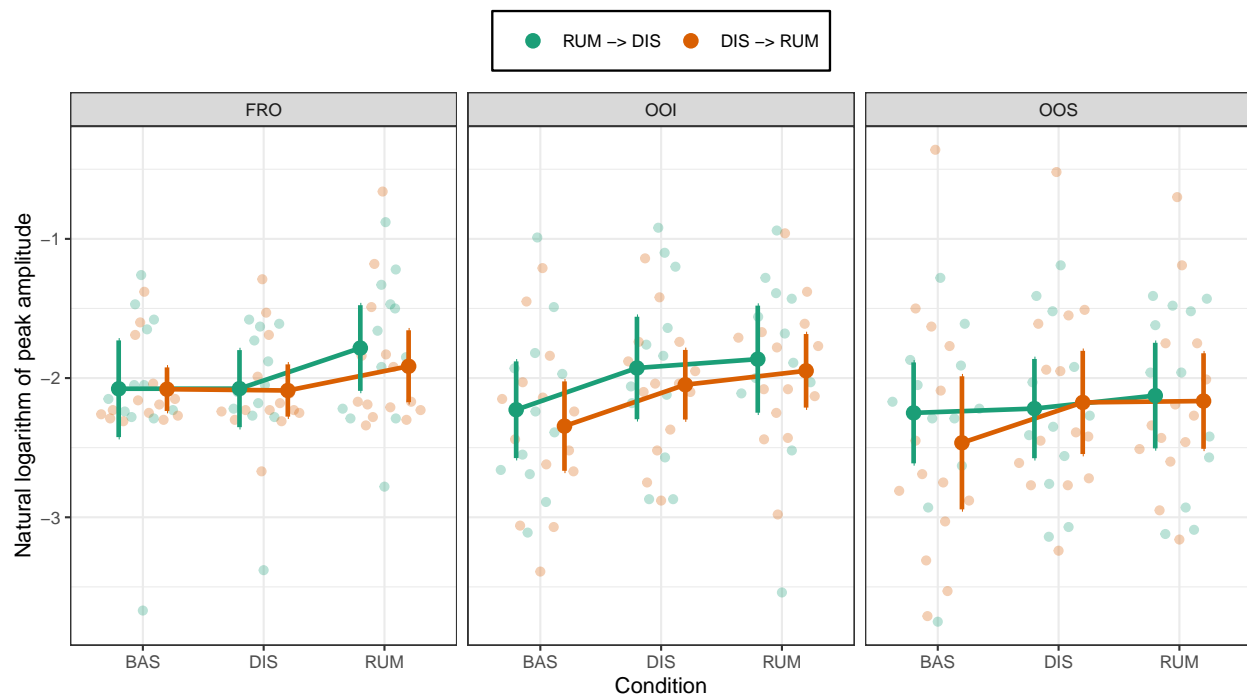

Figure 4. Average log-EMG amplitude by muscle and condition. The black dots and intervals represent the by-group average and 95% confidence interval (N = 26). The horizontal white line in the violin plot represents the median. The grey dots represent the individual-level average natural logarithm of the EMG amplitude by muscle and condition.

About the order effects, Moffatt et al. (2020) say:

"Unless otherwise reported, the inclusion of order in which the conditions were completed as a between-subjects variable as part of a mixed-design ANOVA produced no significant main effects or interactions involving order."

203   Unfortunately, the same line of reasoning applies for testing the effect of the order,

204 which is even less powered than the test of the main effect of interest, rendering it

205 practically uninformative...

Does everyone show the effect?

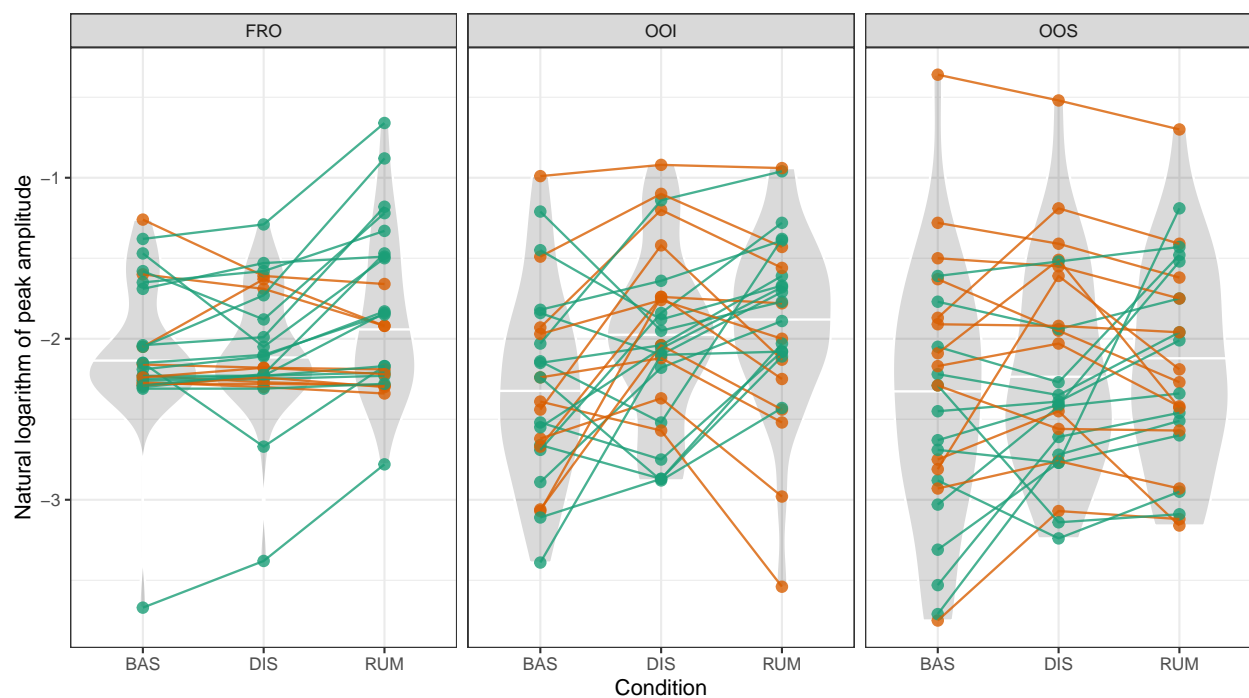207   Haaf and Rouder (2017)...

Figure 5. Inter-individual variability in the main effect of interest (i.e., the difference between the rumination and distraction conditions). Green dots and lines represent the average natural logarithm of the EMG amplitude of participants that showed a higher EMG amplitude in the rumination condition than in the distraction condition, whereas orange dots and lines represent the average natural logarithm of the EMG amplitude of participants that showed a higher EMG amplitude in the distraction condition than in the rumination condition.

208   Huge inter-individual variability... which leads to the next point, what is the relation

209 between self-reports and EMG?

210 **Relating the subjective inner experience to the psychophysiological correlates**

211   ...

Discussion and conclusions

213   ...

Supplementary materials

215   Reproducible code and figures are available at

216 https://github.com/lnalborczyk/inner_experience_EMG.

Acknowledgements

218   Acknowledgements will be included in the final version of this manuscript.

References

220 Alderson-Day, B., & Fernyhough, C. (2015). Inner speech: Development, cognitive

221   functions, phenomenology, and neurobiology. *Psychological Bulletin*, *141*(5), 931–965.

222   https://doi.org/10.1037/bul0000021

223 Aust, F., & Barth, M. (2017). *papaja: Create APA manuscripts with R Markdown.*

224   https://github.com/crsh/papaja

225 Bürkner, P.-C. (2017). brms: An R package for bayesian multilevel models using Stan.

226   *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

227 Ehring, T., & Watkins, E. R. (2008). Repetitive negative thinking as a transdiagnostic

228   process. *International Journal of Cognitive Therapy, 1*(3), 192–205.

229   https://doi.org/10.1680/ijct.2008.1.3.192

230 Goldwin, M., & Behar, E. (2012). Concreteness of idiographic periods of worry and

231   depressive rumination. *Cognitive Therapy and Research, 36*(6), 840–846.

232   https://doi.org/10.1007/s10608-011-9428-1

233 Goldwin, M., Behar, E., & Sibrava, N. J. (2013). Concreteness of depressive rumination

234   and trauma recall in individuals with elevated trait rumination and/or posttraumatic

235   stress symptoms. *Cognitive Therapy and Research, 37*(4), 680–689.

236   https://doi.org/10.1007/s10608-012-9507-y

237 Haaf, J. M., & Rouder, J. N. (2017). Developing constraint in Bayesian mixed models.

238   *Psychological Methods, 22*(4), 779–798. https://doi.org/10.1037/met0000156

239 Lakens, D. (2014). The 20% Statistician: Observed power, and what to do if your editor

240   asks for post-hoc power analyses. In *The 20% Statistician.*

241 Lœvenbruck, H., Grandchamp, R., Rapin, L., Nalborczyk, L., Dohen, M., Perrier, P.,

242   Baciu, M., & Perrone-Bertolotti, M. (2018). A cognitive neuroscience view of inner

243   language: To predict and to hear, see, feel. In P. Langland-Hassan & A. Vicente (Eds.),

244   *Inner speech: New voices* (p. 37). Oxford University Press.

245 Marwick, B. (2019). *Wordcountaddin: Word counts and readability statistics in r

246   markdown documents.*

247 Mayo, D. G. (2018). *Statistical Inference as Severe Testing: How to Get Beyond the

248   Statistics Wars.* Cambridge University Press. https://doi.org/10.1017/9781107286184

249 Mayo, D. G., & Spanos, A. (2006). Severe Testing as a Basic Concept in a NeymanPearson

250 Philosophy of Induction. *The British Journal for the Philosophy of Science*, *57*(2),

251 323–357. https://doi.org/10.1093/bjps/axl003

252 McLaughlin, K. A., Borkovec, T. D., & Sibrava, N. J. (2007). The effects of worry and

253 rumination on affect states and cognitive activity. *Behavior Therapy*, *38*(1), 23–38.

254 https://doi.org/10.1016/j.beth.2006.03.003

255 Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the

256 slow progress of soft psychology. *Journal of Consulting and Clinical Psychology*, *46*(4),

257 806–834. https://doi.org/10.1037/0022-006X.46.4.806

258 Meehl, P. E. (1990a). Why Summaries of Research on Psychological Theories are Often

259 Uninterpretable. *Psychological Reports.* https://doi.org/10.2466/pr0.1990.66.1.195

260 Meehl, P. E. (1997). The problem is epistemology, not statistics: Replace significance tests

261 by confidence intervals and quantify accuracy of risky numerical predictions. *What If*

262 *There Were No Significance Tests?*, 393–425.

263 Meehl, P. E. (1990b). Appraising and Amending Theories: The Strategy of Lakatosian

264 Defense and Two Principles that Warrant It. *Psychological Inquiry*, *1*(2), 108–141.

265 https://doi.org/10.1207/s15327965pli0102_1

266 Meehl, P. E. (1967). Theory-testing in Psychology and Physics: A methodological paradox.

267 *Philosophy of Science*, *34*(2), 103–115. https://doi.org/10.1086/288135

268 Moffatt, J., Mitrenga, K. J., Alderson-Day, B., Moseley, P., & Fernyhough, C. (2020).

269 Inner experience differs in rumination and distraction without a change in

270 electromyographical correlates of inner speech. *PLOS ONE*, *15*(9), e0238920.

271 https://doi.org/10.1371/journal.pone.0238920

272 Morey, R. D., & Rouder, J. N. (2018). *BayesFactor: Computation of bayes factors for*

273 *common designs.* https://CRAN.R-project.org/package=BayesFactor

274 Müller, K. (2017). *Here: A simpler way to find your files.*

275 https://CRAN.R-project.org/package=here

276 Nalborczyk, L. (2019). *Understanding rumination as a form of inner speech: Probing the*

277 *role of motor processes* [PhD Thesis]. Univ. Grenoble Alpes & Ghent University.

278 Nalborczyk, L., Banjac, S., Celine, B., Grandchamp, R., Koster, E. H. W., Marcela, P.-B.,

279 & Loevenbruck, H. (2020). *Dissociating facial electromyographic correlates of visual and*

280 *verbal induced rumination.* https://doi.org/10.31234/osf.io/vfjn2

281 Nalborczyk, L., Grandchamp, R., Koster, E. H. W., Perrone-Bertolotti, M., & Lœvenbruck,

282 H. (2020). Can we decode phonetic features in inner speech using surface

283 electromyography? *PLOS ONE, 15*(5), e0233282.

284 https://doi.org/10.1371/journal.pone.0233282

285 Nalborczyk, L., Perrone-Bertolotti, M., Baeyens, C., Grandchamp, R., Polosan, M.,

286 Spinelli, E., Koster, E. H. W., & Lœvenbruck, H. (2017). Orofacial electromyographic

287 correlates of induced verbal rumination. *Biological Psychology, 127*, 53–63.

288 https://doi.org/10.1016/j.biopsycho.2017.04.013

289 Nalborczyk, L., Perrone-Bertolotti, M., Baeyens, C., Grandchamp, R., Spinelli, E., Koster,

290 E. H. W., & Lœvenbruck, H. (2020). *Articulatory suppression effects on induced*

291 *rumination* [Under Review].

292 Perrone-Bertolotti, M., Rapin, L., Lachaux, J. P., Baciu, M., & Lœvenbruck, H. (2014).

293 What is that little voice inside my head? Inner speech phenomenology, its role in

294 cognitive performance, and its relation to self-monitoring. *Behavioural Brain Research,*

295 *261*, 220–239. https://doi.org/10.1016/j.bbr.2013.12.034

296 Pollard, P., & Richardson, J. T. (1987). On the probability of making Type I errors.

297   *Psychological Bulletin, 102*(1), 159–163. https://doi.org/10.1037/0033-2909.102.1.159

298 R Core Team. (2017). *R: A language and environment for statistical computing.* R

299   Foundation for Statistical Computing. https://www.R-project.org/

300 Rouder, J. N., Morey, R. D., Verhagen, J., Province, J. M., & Wagenmakers, E.-J. (2016).

301   Is There a Free Lunch in Inference? *Topics in Cognitive Science, 8*(3), 520–547.

302   https://doi.org/10.1111/tops.12214

303 Wickham, H. (2017). *Tidyverse: Easily install and load the 'tidyverse'.*

304   https://CRAN.R-project.org/package=tidyverse

305 Xie, Y. (2015). *Dynamic documents with R and knitr* (2nd ed.). Chapman; Hall/CRC.

306   https://yihui.org/knitr/

307 Xie, Y., Allaire, J. J., & Grolemund, G. (2018). *R markdown: The definitive guide.*

308   Chapman; Hall/CRC. https://bookdown.org/yihui/rmarkdown