CODECHECK certificate 2025-005

https://doi.org/10.5281/zenodo.15323262

Item	Value
Title	Verification Report of Zhao et al. (2023) 'Effect of Acceptance and Commitment Therapy for depressive disorders: a meta-analysis'
Authors	Ian Hussey , Malte Elson
Reference	https://doi.org/10.71240/lcyc.052239
Codechecker	Daniel Nüst
Date of check	2025-05-15 12:00:00
Summary	TODO summary in codecheck.yml
Repository	https://github.com/codecheckers/lifecycle-journal-codechecks

Table 1: CODECHECK summary

Output	Comment	Size (b)
`dataandcode/communciations/plots/verification_zhao_2023_comparisons_plot.pdf`	article Figure 2, generated with notebook	6921
`dataandcode/code/analysis/verification_zhao_2023.html`	computational notebook with code for Figure 1, Figure 2, and Table 3	2229354

Table 2: Summary of output files generated

Summary

The article referenced data and code in a OSF project. An included computational notebook in R Markdown format could be executed without errors and Figure 1, Figure 2, Figure 3, and Table 2 as well as numerical statics could be recreated with the code in the notebook. The data deposit contains further files matching other tables in the article, though these are likely, at least partially, manually created and out of scope. The check was successful.

CODECHECKER notes

This check is based on a submission to the Lifecycle Journal. Basis for evaluation is the registration of the code and data repository linked from the article page at https://doi.org/10.17605/OSF.IO/BMQCW (the registration of the project linked in the article). The used article was published in the Lifecycle Journal Apr 29, 2025 with "Version 1 | Status: Under Evaluation". The materials of the article include data and code, both are published on OSF in the project https://osf.io/jwcmu/. This check starts with the article in PDF form and the archive `data and code.zip` (https://osf.io/pexfb, from the registration).

The article does point to the OSF project several times, but there are no details or instructions. The archive contains a `README.md` file, which however only includes the abstract and the license (CC-BY 4.0). The archive contains three folders with several subfolders. Here is the full file tree:

```
## ../data and code/
## +-- LICENSE
## +-- README.md
## +-- code
## |   \-- analysis
## |       +-- verification_zhao_2023.Rmd
## |       \-- verification_zhao_2023.html
## +-- communciations
## |   +-- plots
## |   |   +-- new_meta.png
## |   |   +-- verification_zhao_2023_comparisons_plot.pdf
## |   |   +-- verification_zhao_2023_comparisons_plot.png
## |   |   +-- verification_zhao_2023_comparisons_plot_alt.pdf
## |   |   \-- verification_zhao_2023_comparisons_plot_alt.png
## |   \-- submission
## |       \-- cover letter.docx
## \-- data
##     +-- processed
##     |   +-- influence_recalculated.xlsx
##     |   +-- influence_zhao.xlsx
##     |   \-- summary_statistics_comparisons.xlsx
##     +-- raw
##     |   \-- data_zhao_forest_plot_1_extracted.xlsx
##     \-- screenshots
##         \-- zhao et al figure 4.png
```

I did a manual matching of the files in the archive and the figures in the article due to a lack of documentation.

Article	File	Visual/manual inspection, comment
Figure 1	-	no corresponding file, but computational notebook section "Sensitivity analysis: Exclude Tjak et al. (2018)" creates a matching the figure
Table 1	`data and code/data/processed/summary_statistics_comparisons.xlsx`	the same, though spreadsheet has extra highlights
Figure 2	`data and code/communciations/plots/verification_zhao_2023_comparisons_plot.pdf/png`	the same
Table 2	`data and code/data/raw/data_zhao_forest_plot_1_extracted.xlsx`	matches a combination of sheets "original" and "reextracted"
Table 3	-	no corresponding file, matches the table in notebook section "Outlier/undue influence tests"
Figure 3	`data and code/communciations/plots/new_meta.png`	also matches notebook section "Sensitivity analysis: Exclude Zemestani et al. (2020)", it is unclear why this is not also saved as a plot in the notebook
-	`data and code/data/processed/influence_recalculated.xlsx`	no matching table found
-	`data and code/data/processed/influence_zhao.xlsx`	no matching table found, but matches table generated in notebook section "Outlier/undue influence tests"

Article	File	Visual/manual inspection, comment
-	`data and code/communciations/plots/verification_zhao_2023_comparisons_plot_alt.pdf/png`	mo matching figure found

I add the matching files to the manifest for this check (see `codecheck.yml`). Next, I check out the computational notebook in `data and code/code/analysis/`, which is an R Markdown file, `verification_zhao_2023.Rmd` that is also provided in a rendered output file, `verification_zhao_2023.html`. Based on the provided information, I expand the table above.

I continue with the `.Rmd` file, beginning with installing missing packages, which could be installed without any errors. Here is the output of my `sessionInfo()` after successfully rendering the computational notebook using RStudio's "Knit"-button without any errors within a few moments on my laptop:

```
R version 4.5.0 (2025-04-11)
Platform: x86_64-pc-linux-gnu
Running under: Ubuntu 22.04.5 LTS

Matrix products: default
BLAS:   /usr/lib/x86_64-linux-gnu/blas/libblas.so.3.10.0
LAPACK: /usr/lib/x86_64-linux-gnu/lapack/liblapack.so.3.10.0  LAPACK version 3.10.0

locale:
 [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C               LC_TIME=de_DE.UTF-8        LC_COLLATE=en_US.UTF-8     LC_MONETARY=de_DE.UTF-8    LC_MESSAGES=en_US.UTF-8    LC_PAPER=de_DE.UTF-8
 [8] LC_NAME=C                  LC_ADDRESS=C               LC_TELEPHONE=C             LC_MEASUREMENT=de_DE.UTF-8 LC_IDENTIFICATION=C

time zone: Europe/Berlin
tzcode source: system (glibc)

attached base packages:
[1] stats     graphics  grDevices datasets  utils     methods   base

other attached packages:
 [1] lubridate_1.9.4  forcats_1.0.0    stringr_1.5.1    dplyr_1.1.4      purrr_1.0.4      tidyr_1.3.1      ggplot2_3.5.2    tidyverse_2.0.0  readr_2.1.5      tibble_3.2.1     xtable_1.8-4
[12] yaml_2.3.10      rprojroot_2.0.4  knitr_1.50       codecheck_0.14.0 parsedate_1.3.2  R.cache_0.16.0   gh_1.4.1

loaded via a namespace (and not attached):
 [1] tidyselect_1.2.1  R.utils_2.13.0    fastmap_1.2.0     XML_3.99-0.18     digest_0.6.37     timechange_0.3.0  lifecycle_1.0.4   qpdf_1.3.5        magrittr_2.0.3    zen4R_0.10        compiler_4.5.0
[12] rlang_1.1.6       sass_0.4.10       tools_4.5.0       utf8_1.2.4        fauxpas_0.5.2     askpass_1.2.1     bit_4.6.0         curl_6.2.2        plyr_1.8.9        xml2_1.3.8        pkgload_1.4.0
[23] redland_1.0.17-18 httpcode_0.3.0    withr_3.0.2       rdflib_0.2.9      R.oo_1.27.0       grid_4.5.0        roxygen2_7.3.2    atom4R_0.3-3      rorcid_0.7.0      colorspace_2.1-1  scales_1.3.0
[34] crul_1.5.0        cli_3.6.5         rmarkdown_2.29    crayon_1.5.3      generics_0.1.3    rstudioapi_0.17.1 httr_1.4.7        tzdb_0.5.0        dadjokeapi_1.0.2  cachem_1.1.0      assertthat_0.2.1
[45] parallel_4.5.0    vctrs_0.6.5       jsonlite_2.0.0    hms_1.1.3         bit64_4.6.0-1     jquerylib_0.1.4   keyring_1.3.2     glue_1.8.0        stringi_1.8.7     gtable_0.3.6      munsell_0.5.1
[56] bspm_0.5.7        pillar_1.10.2     htmltools_0.5.8.1 R6_2.6.1          vroom_1.6.5       evaluate_1.0.3    osfr_0.2.9        R.methodsS3_1.8.2 png_0.1-8         memoise_2.0.1     bslib_0.9.0
[67] Rcpp_1.0.14       zip_2.3.2         whisker_0.4.1     xfun_0.52         fs_1.6.6          pdftools_3.5.0    pkgconfig_2.0.3
```

Further statistics calculated in the computational notebook match statistical numbers in the article, though a precise check is out of scope for this evaluation due to missing documentation.

Recommendations

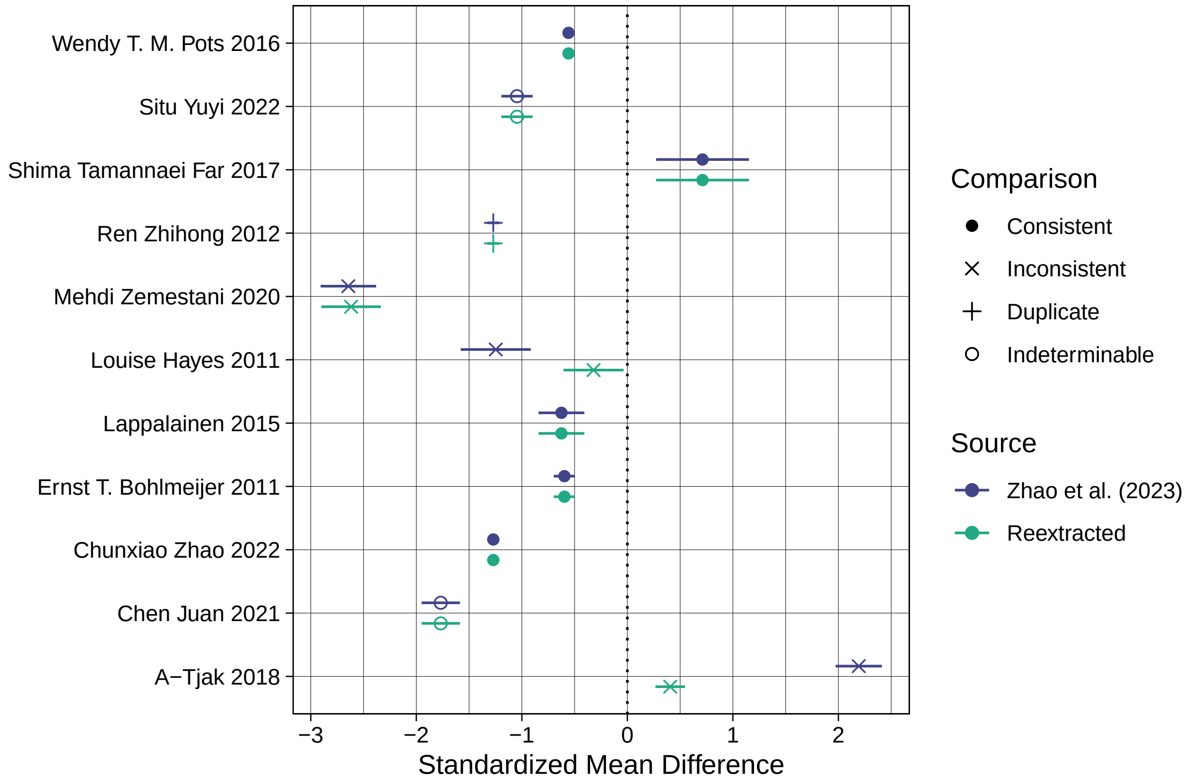
I suggest to the authors to consider the following suggestions for their next publication or workflow:

- Use appropriate data and software licenses - CC-BY is not ideal for either.
- Use precise file names in your data export and expand README to make clear which file corresponds to which figure; all files should be saved through a code statement (not done for `new_meta.png`).
- Provide documentation or even snapshot of the used packages (`sessionInfo()`, `renv` project, ..).
- Very clearly mention in the article where in the computational notebook statistics are calculated.
- Reference the registration in your article, ideally as a proper citation.

Manifest files

verification_zhao_2023_comparisons_plot.pdf

Comment: article Figure 2, generated with notebook



verification_zhao_2023.html

Content of HTML file (starts on next page):

Reproduce Zhao’s results
Re-extracted
Preemptive response

Verification of Zhao et al. (2023)

Code

Ian Hussey

Show

Reproduce Zhao’s results

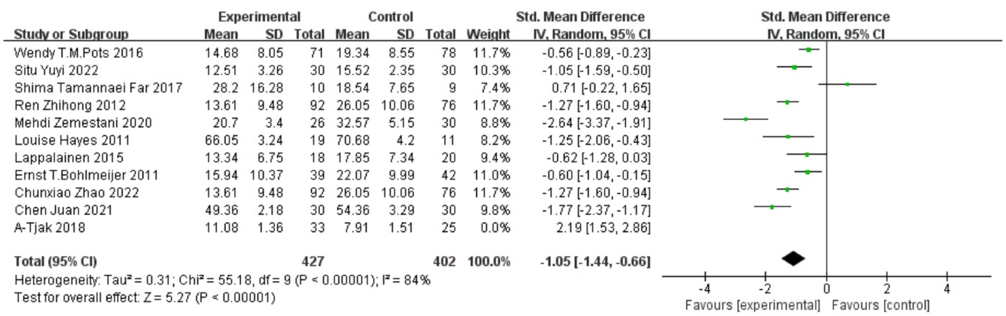
Zhao et al. (2023) present the results of multiple meta-analyses, sensitivity analyses, and subgroup analyses. Here, I attempt to reproduce and verify the result of only the first meta-analysis reported in the first forest plot (figure 4), which included the largest number of studies and data for the original studies’ post-intervention time point.

Zhao et al. (2023, p. 6) reported “Effect of ACT on depression: Eleven studies reported the effect of ACT on depression levels in 887 patients with depressive disorders. There was statistically significant heterogeneity among the studies (P < 0.00001; I2= 93%), therefore, the random effects model was used to conduct meta-analysis. There was significant treatment effect from ACT in comparison to other treatments, as shown in the forest plot in Fig. 4.”

However, Figure 4 does not present the full meta-analysis, but rather a sensitivity analysis after excluding A-Tjak et al. (2018). The rationale for excluding that study is not stated. I therefore first fit a meta-analysis to all 11 studies, then perform a series of outlier/undue influence tests to try to reproduce the conclusion that A-Tjak et al. (2018) should be excluded, and then finally I fit a new meta-analysis to the remaining 10 studies to try to reproduce the meta-analytic effect reported in Zhao et al. (2023) Figure 4.

Summary stats (M, SD, N for both groups) were extracted from Zhao et al.’s (2023) forest plot in Figure 4. Specifically:

Show



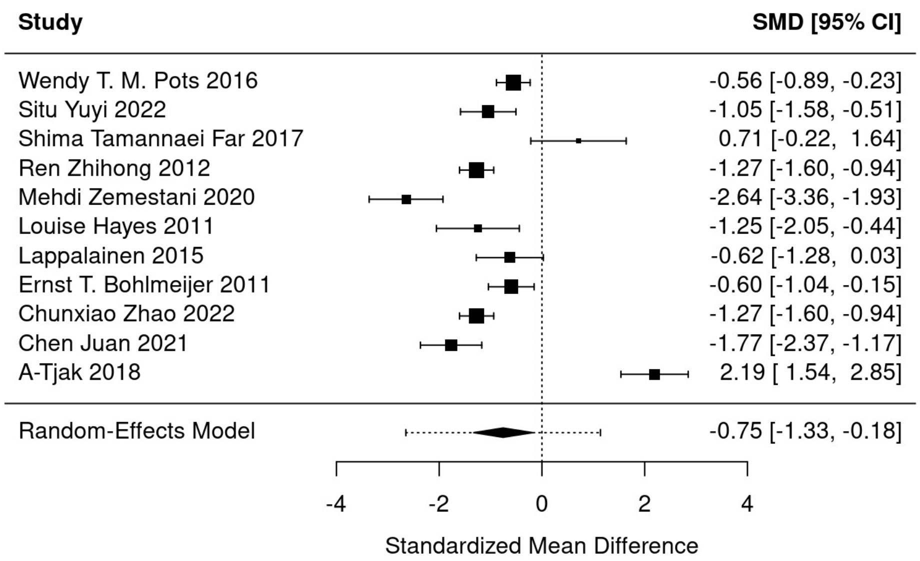
	Experimental			Control				Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
Wendy T.M.Pots 2016	14.68	8.05	71	19.34	8.55	78	11.7%	-0.56 [-0.89, -0.23]
Situ Yuyi 2022	12.51	3.26	30	15.52	2.35	30	10.3%	-1.05 [-1.59, -0.50]
Shima Tamannaei Far 2017	28.2	16.28	10	18.54	7.65	9	7.4%	0.71 [-0.22, 1.65]
Ren Zhihong 2012	13.61	9.48	92	26.05	10.06	76	11.7%	-1.27 [-1.60, -0.94]
Mehdi Zemestani 2020	20.7	3.4	26	32.57	5.15	30	8.8%	-2.64 [-3.37, -1.91]
Louise Hayes 2011	66.05	3.24	19	70.68	4.2	11	8.2%	-1.25 [-2.06, -0.43]
Lappalainen 2015	13.34	6.75	18	17.85	7.34	20	9.4%	-0.62 [-1.28, 0.03]
Ernst T.Bohlmeijer 2011	15.94	10.37	39	22.07	9.99	42	11.0%	-0.60 [-1.04, -0.15]
Chunxiao Zhao 2022	13.61	9.48	92	26.05	10.06	76	11.7%	-1.27 [-1.60, -0.94]
Chen Juan 2021	49.36	2.18	30	54.36	3.29	30	9.8%	-1.77 [-2.37, -1.17]
A-Tjak 2018	11.08	1.36	33	7.91	1.51	25	0.0%	2.19 [1.53, 2.86]
Total (95% CI)			**427**			**402**	**100.0%**	**-1.05 [-1.44, -0.66]**

Heterogeneity: $Tau^2 = 0.31$; $Chi^2 = 55.18$, df = 9 (P < 0.00001); $I^2 = 84\%$
Test for overall effect: Z = 5.27 (P < 0.00001)

Fig. 4 Effect of ACT on depression in patients with depressive disorders after sensitivity analysis

Meta-analysis of all 11 studies

Show



Show

```
##
## Random-Effects Model (k = 11; tau^2 estimator: DL)
##
## tau^2 (estimated amount of total heterogeneity): 0.8485 (SE = 0.4826)
## tau (square root of estimated tau^2 value):      0.9212
## I^2 (total heterogeneity / total variability):   93.11%
## H^2 (total variability / sampling variability):  14.51
##
## Test for Heterogeneity:
## Q(df = 10) = 145.0800, p-val < .0001
##
## Model Results:
##
## estimate      se     zval    pval    ci.lb    ci.ub
##  -0.7548  0.2925  -2.5809  0.0099  -1.3280  -0.1816  **
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Heterogeneity: Q(df = 10) = 145.08, p < .0001, tau^2 = 0.8485, I^2 = 93.1%, H^2 = 14.51

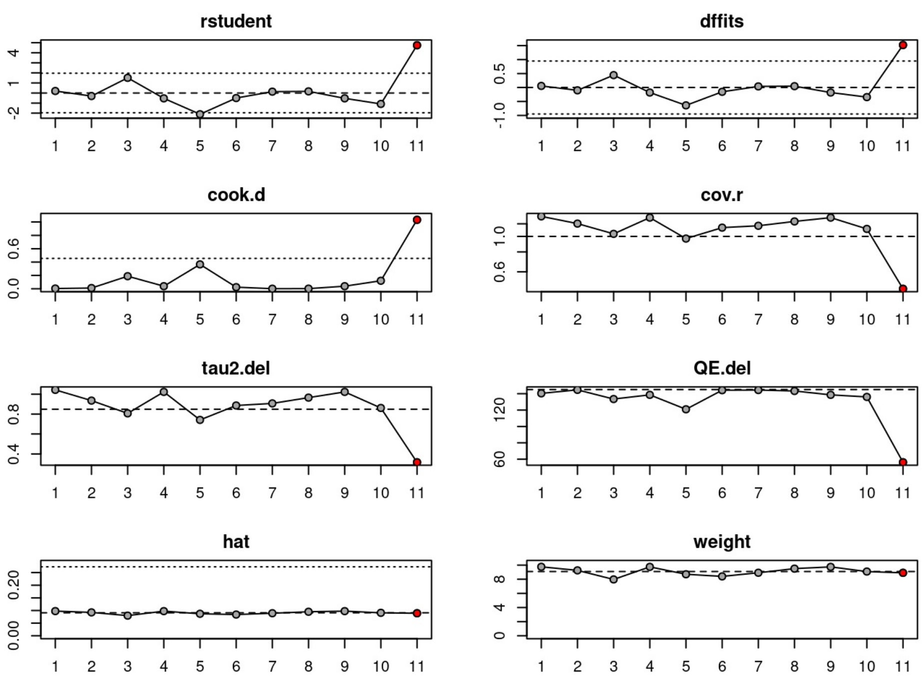
Outlier/undue influence tests

No specific justification was given in Zhao et al. (2023) for their exclusion of A-Tjak et al. (2018). In the absence of this, I employ {metafor}'s `influence()` function which applies multiple different metrics of undue influence.

Show

```
##
##                            rstudent dffits cook.d cov.r tau2.del QE.del  hat
## Wendy T. M. Pots 2016          0.20   0.06   0.00  1.34     1.04 140.42 0.10
## Situ Yuyi 2022                -0.30  -0.10   0.01  1.20     0.94 144.70 0.09
## Shima Tamannaei Far 2017       1.51   0.44   0.19  1.04     0.81 133.49 0.08
## Ren Zhihong 2012              -0.53  -0.18   0.04  1.31     1.02 138.61 0.10
## Mehdi Zemestani 2020          -2.11  -0.64   0.37  0.97     0.74 120.95 0.09
## Louise Hayes 2011             -0.50  -0.15   0.02  1.14     0.89 144.26 0.08
## Lappalainen 2015               0.13   0.04   0.00  1.17     0.91 144.45 0.09
## Ernst T. Bohlmeijer 2011       0.16   0.05   0.00  1.24     0.97 143.33 0.10
## Chunxiao Zhao 2022            -0.53  -0.18   0.04  1.31     1.02 138.61 0.10
## Chen Juan 2021                -1.09  -0.34   0.12  1.12     0.86 136.07 0.09
## A-Tjak 2018                    4.74   1.52   1.03  0.47     0.32  56.20 0.09
##                            weight  dfbs inf
## Wendy T. M. Pots 2016        9.76  0.06
## Situ Yuyi 2022               9.25 -0.10
## Shima Tamannaei Far 2017     7.97  0.44
## Ren Zhihong 2012             9.75 -0.18
## Mehdi Zemestani 2020         8.70 -0.64
## Louise Hayes 2011            8.40 -0.15
## Lappalainen 2015             8.92  0.04
## Ernst T. Bohlmeijer 2011     9.50  0.05
## Chunxiao Zhao 2022           9.75 -0.18
## Chen Juan 2021               9.09 -0.34
## A-Tjak 2018                  8.91  1.55   *
```

Show

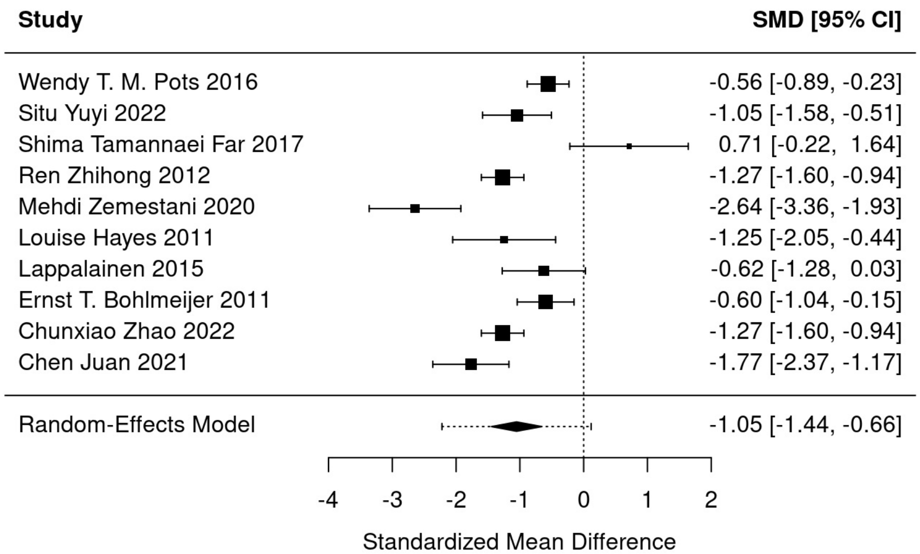


A-Tjak was detected as an outlier. Although Zhao et al.'s (2023) specific method of outlier detection is unknown, their conclusion is replicate by these results.

Sensitivity analysis: Exclude Tjak et al. (2018)

As in Zhao Figure 4, and following the results of the above outlier tests, I exclude A-Tjak et al. (2018) and fit a new meta-analysis in order to attempt to reproduce the results reported in Figure 4.

Show



Show

```
##
## Random-Effects Model (k = 10; tau^2 estimator: DL)
##
## tau^2 (estimated amount of total heterogeneity): 0.3159 (SE = 0.2036)
## tau (square root of estimated tau^2 value):      0.5621
## I^2 (total heterogeneity / total variability):   83.99%
## H^2 (total variability / sampling variability):  6.24
##
## Test for Heterogeneity:
## Q(df = 9) = 56.2016, p-val < .0001
##
## Model Results:
##
## estimate      se     zval    pval    ci.lb    ci.ub
##  -1.0520  0.2004  -5.2504  <.0001  -1.4447  -0.6593  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Zhao et al.'s (2023) forest plot in Figure 4 reported: Meta-analytic d = -1.05, 95% CI [-1.44, -0.66]. The results above therefore reproduce this result from the summary statistics reported in that forest plot.

Heterogeneity: Q(df = 9) = 56.2, p < .0001, tau^2 = 0.3159, I^2 = 84%, H^2 = 6.24

NNT

NNT = 1.84, 95% CI [1.44, 2.79]

Re-extracted

Although the results reported in Figure 4 can be reproduced from the summary statistics reported in Figure 4, the validity of those results are reliant on the summary statistics being accurately extracted from the original studies. Recent work has shown that this is frequently not the case (e.g., Maassen et al., 2020). I therefore attempted to extract the Mean, SD, and N for the ACT and control conditions at the post-intervention time point for each of the original studies.

In the below plots, I use the same labelling as in Zhao et al. (2023). The references (including DOIs) for the studies they refer to are available in the excel files that the data is loaded from. Note that there is imperfect correspondence between Zhao's labels and the authors of the original studies.

Comparisons of extracted summary stats

Show

comparison	n
Consistent	5
Duplicate	1
Inconsistent	3
Indeterminable	2

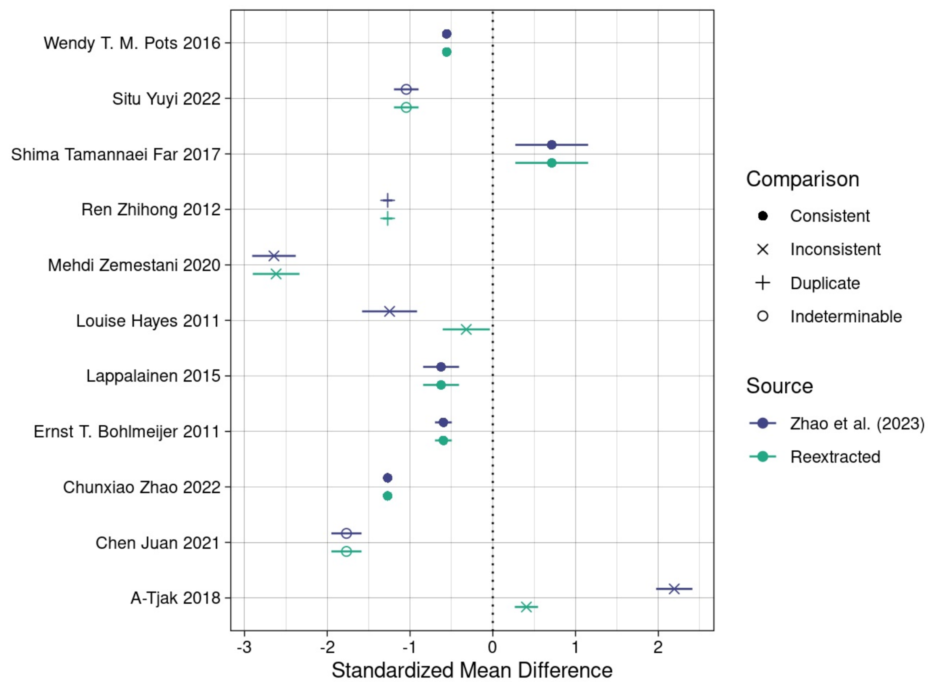
Two articles with DOIs were only available in Chinese for a fee, and were therefore not accessed. One article did not have a DOI listed and could not be found via internet searches. It is possible that the reference listed in Zhao et al. (2023) is an English translation of the original Chinese title. As such, the reproducibility of 3 of the 11 studies was indeterminable.

In 3 of the remaining 8 articles, the summary statistics (M, SD, N) reported in Zhao et al. (2023, Figure 4) do not match those reported in the original studies. In two cases (Louise Hayes 2011, A-Tjak 2018), the original articles reported Standard Errors and these were employed by Zhao et al. (2023) as if they were Standard Deviations. This is a common and unfortunate issue in meta analyses, as is problematic as SEs are much smaller than SDs, therefore inflating the apparent Standardized Mean Difference effect sizes employed in the meta-analaysis. In one case (Mehdi Zemestani 2020), the Ns do not match those reported in the original publication, although the nature of the error is unclear. As such, only 5 of the 11 sets of summary statistics could be verified against the numbers reported in the original articles.

Show

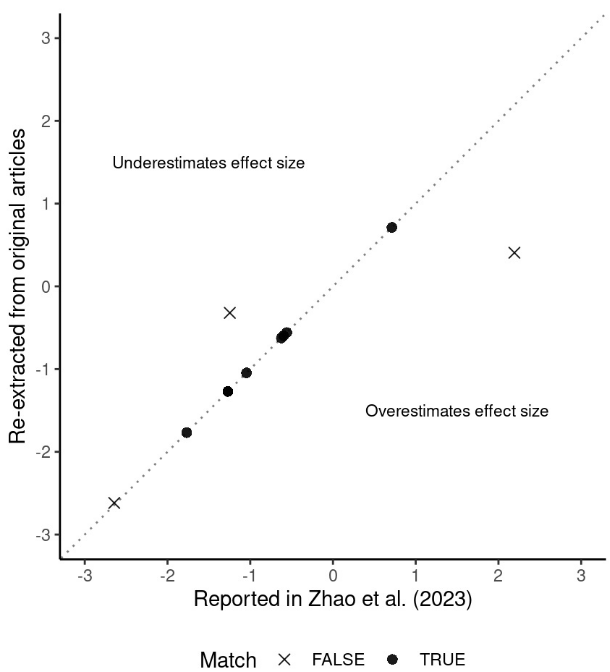
study	source	mean_exp	sd_exp	n_exp	mean_control	sd_control	n_control	reextracted	comparison
A-Tjak 2018	Zhao et al. (2023)	11.08	1.36	33	7.91	1.51	25	NA	
A-Tjak 2018	Reextracted	11.08	7.81	33	7.91	7.55	25	TRUE	Inconsistent
Chen Juan 2021	Zhao et al. (2023)	49.36	2.18	30	54.36	3.29	30	NA	
Chen Juan 2021	Reextracted	49.36	2.18	30	54.36	3.29	30	FALSE	Indeterminable
Chunxiao Zhao 2022	Zhao et al. (2023)	13.61	9.48	92	26.05	10.06	76	NA	
Chunxiao Zhao 2022	Reextracted	13.61	9.48	92	26.05	10.06	76	TRUE	Consistent
Ernst T. Bohlmeijer 2011	Zhao et al. (2023)	15.94	10.37	39	22.07	9.99	42	NA	
Ernst T. Bohlmeijer 2011	Reextracted	15.94	10.37	39	22.07	9.99	42	TRUE	Consistent
Lappalainen 2015	Zhao et al. (2023)	13.34	6.75	18	17.85	7.34	20	NA	
Lappalainen 2015	Reextracted	13.34	6.75	18	17.85	7.34	20	TRUE	Consistent
Louise Hayes 2011	Zhao et al. (2023)	66.05	3.24	19	70.68	4.20	11	NA	
Louise Hayes 2011	Reextracted	66.05	14.12	19	70.68	13.93	11	TRUE	Inconsistent
Mehdi Zemestani 2020	Zhao et al. (2023)	20.70	3.40	26	32.57	5.15	30	NA	
Mehdi Zemestani 2020	Reextracted	20.70	3.40	23	32.57	5.15	29	TRUE	Inconsistent
Ren Zhihong 2012	Zhao et al. (2023)	13.61	9.48	92	26.05	10.06	76	NA	
Ren Zhihong 2012	Reextracted	13.61	9.48	92	26.05	10.06	76	FALSE	Duplicate
Shima Tamannaei Far 2017	Zhao et al. (2023)	28.20	16.28	10	18.54	7.65	9	NA	
Shima Tamannaei Far 2017	Reextracted	28.20	16.28	10	18.54	7.65	9	TRUE	Consistent
Situ Yuyi 2022	Zhao et al. (2023)	12.51	3.26	30	15.52	2.35	30	NA	
Situ Yuyi 2022	Reextracted	12.51	3.26	30	15.52	2.35	30	FALSE	Indeterminable
Wendy T. M. Pots 2016	Zhao et al. (2023)	14.68	8.05	71	19.34	8.55	78	NA	
Wendy T. M.	Reextracted	14.68	8.05	71	19.34	8.55	78	TRUE	Consistent

Show



Show

Show

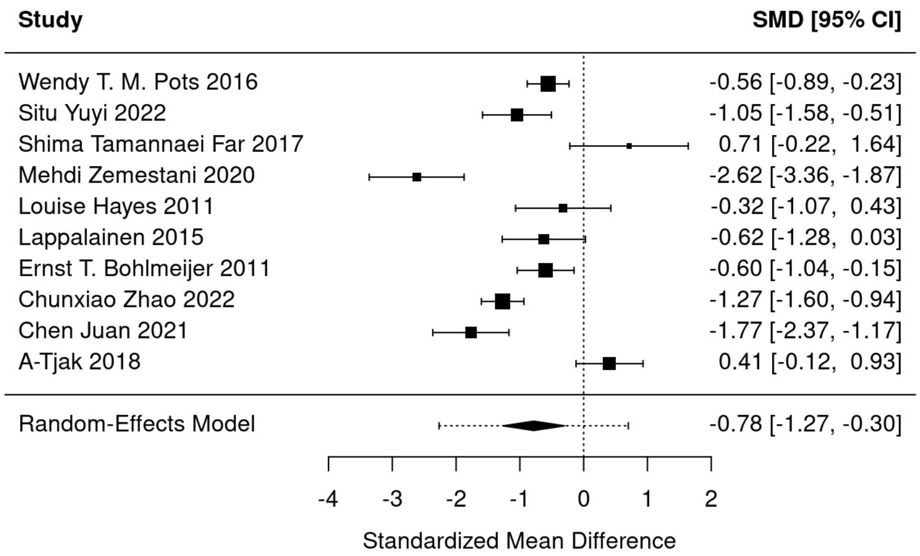


Show

Updated meta-analysis of the 10 studies after excluding the duplicate

Following the same workflow as Zhao et al. (2023), a random effects meta-analysis was then fit to the 10 sets of effect sizes (i.e., including the 3 that were indeterminable, on the liberal assumption that they are correct, but excluding the duplicate on the reasonable assumption that it is not).

Show



Show

```
##
## Random-Effects Model (k = 10; tau^2 estimator: DL)
##
## tau^2 (estimated amount of total heterogeneity): 0.5125 (SE = 0.3143)
## tau (square root of estimated tau^2 value):      0.7159
## I^2 (total heterogeneity / total variability):   88.28%
## H^2 (total variability / sampling variability):  8.53
##
## Test for Heterogeneity:
## Q(df = 9) = 76.7813, p-val < .0001
##
## Model Results:
##
## estimate      se     zval    pval    ci.lb    ci.ub
##  -0.7831  0.2459  -3.1848  0.0014  -1.2651  -0.3012  **
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Heterogeneity: Q(df = 9) = 76.78, p < .0001, tau^2 = 0.5125, I^2 = 88.3%, H^2 = 8.53

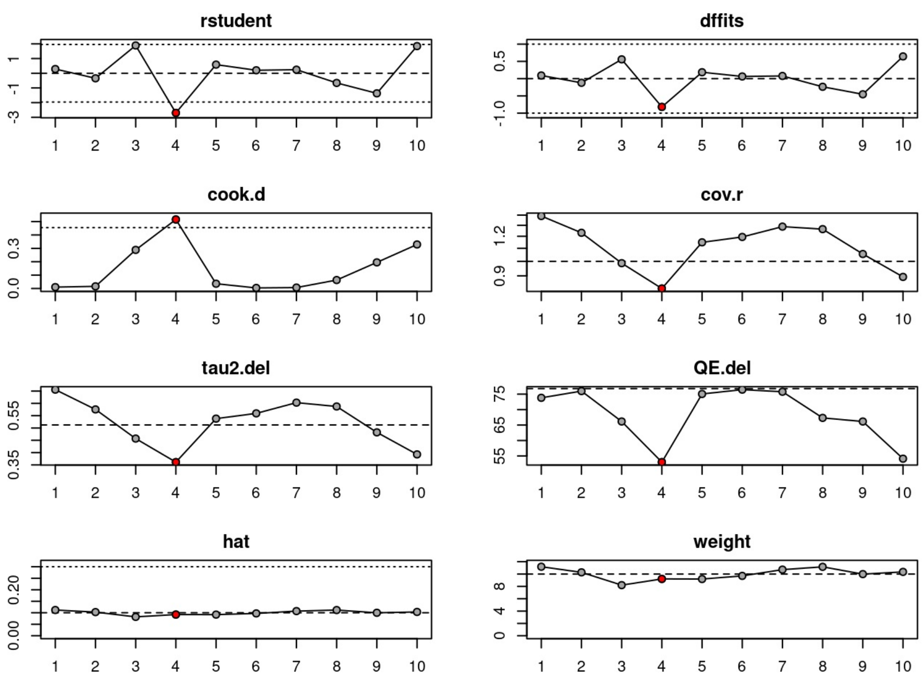
Outlier/undue influence tests

Similarly, I then checked for outliers/undue influence.

Show

```
##
##                            rstudent dffits cook.d cov.r tau2.del QE.del  hat
## Wendy T. M. Pots 2016          0.29   0.09   0.01  1.39     0.66  73.81 0.11
## Situ Yuyi 2022                -0.34  -0.12   0.02  1.23     0.58  75.98 0.10
## Shima Tamannaei Far 2017       1.89   0.56   0.29  0.99     0.46  66.15 0.08
## Mehdi Zemestani 2020          -2.70  -0.82   0.52  0.82     0.36  53.00 0.09
## Louise Hayes 2011              0.59   0.19   0.04  1.15     0.54  75.05 0.09
## Lappalainen 2015               0.20   0.06   0.00  1.19     0.56  76.45 0.10
## Ernst T. Bohlmeijer 2011       0.24   0.08   0.01  1.29     0.60  75.77 0.11
## Chunxiao Zhao 2022            -0.66  -0.24   0.06  1.26     0.59  67.31 0.11
## Chen Juan 2021                -1.37  -0.45   0.20  1.05     0.48  66.14 0.10
## A-Tjak 2018                    1.85   0.64   0.33  0.89     0.39  54.12 0.10
##                            weight  dfbs inf
## Wendy T. M. Pots 2016       11.19  0.09
## Situ Yuyi 2022              10.28 -0.12
## Shima Tamannaei Far 2017     8.20  0.56
## Mehdi Zemestani 2020         9.21 -0.83   *
## Louise Hayes 2011            9.19  0.19
## Lappalainen 2015             9.70  0.06
## Ernst T. Bohlmeijer 2011    10.72  0.08
## Chunxiao Zhao 2022          11.17 -0.24
## Chen Juan 2021               9.99 -0.45
## A-Tjak 2018                 10.35  0.64
```

Show



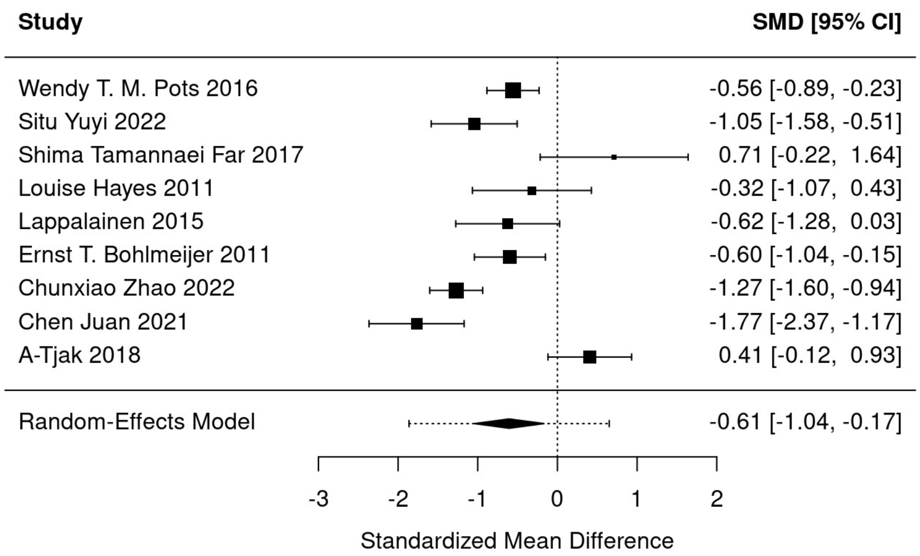
This suggests that Zemestani et al. (2020) is an outlier that should be excluded, whereas Zhao et al. (2023) excluded A-Tjak et al. (2018).

Note that there are additional reasons to exclude the results of Zemestani et al. (2020): as documented in this pubpeer comment (https://pubpeer.com/publications/0E13E34679B18385D6C4C29143A9CD), several results reported in that article fail StatCheck, GRIM, and GRIMMER tests (i.e., the reported results are mathmatically impossible); the reported SD of the BDI-II scores is implausibly small given normative data for the BDI-II elsewhere, and as a result the SMD effect size is implausibly large (d = -2.64 in Zhao et al. 2023).

Sensitivity analysis: Exclude Zemestani et al. (2020)

I therefore fitted a new meta-analysis to the remaining 9 studies after excluding Zemestani et al. (2020).

Show



Show

```
##
## Random-Effects Model (k = 9; tau^2 estimator: DL)
##
## tau^2 (estimated amount of total heterogeneity): 0.3613 (SE = 0.2398)
## tau (square root of estimated tau^2 value):      0.6011
## I^2 (total heterogeneity / total variability):   84.90%
## H^2 (total variability / sampling variability):  6.62
##
## Test for Heterogeneity:
## Q(df = 8) = 52.9976, p-val < .0001
##
## Model Results:
##
## estimate      se     zval    pval    ci.lb    ci.ub
##  -0.6065  0.2227  -2.7230  0.0065  -1.0431  -0.1700  **
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Show

A significant meta-analytic effect was still found, but this effect was -42.3% smaller than that reported by Zhao et al. (2023) after correcting the data extraction errors and refitting the meta-analysis and sensitivity analysis workflow.

Heterogeneity: Q(df = 8) = 53, p < .0001, tau^2 = 0.3613, I^2 = 84.9%, H^2 = 6.62

NNT

NNT = 3.01, 95% CI [1.85, 10.45]

Preemptive response

To preempt one possible response to these concerns: readers might argue something along the lines of "the conclusions of Zhao et al. (2023) are not substantially affected by the correction of these errors, and therefore there is no strong need for concern." However, any such reactions should be seen as worrisome rather than encouraging: meta-analyses are often argued to be at the top of the Evidence Pyramid, and their goal is to precisely estimate the true effect size beyond merely estimating whether the true effect size is non-zero or not. To retreat to the weaker position that the obtained effect size does not substantively matter is to undermine the point of doing the meta-analysis in the first place. Either the empirical estimates matter or they don't.

Finally, as discussed at the start of this document, no attempt was made to reproduce or attempt to understand the impact of the detected extraction errors on the other analyses reported in Zhao et al. (2023). It is likely that the subgroup analyses reported after Figure 4 will also be affected by the issues detected above. It is also possible that other issues exist, given that errors were made at the relatively simple data extraction phase. The results and conclusions of Zhao et al. (2023) should therefore be regarded with great skepticism given the issues already detected, and require careful revision throughout to ensure no other errors were made.

Equally, this verification does not attempt to critique things that were not done by Zhao et al. (2023). For example, the authors do not apply any method of bias correction for p-hacking or publication bias, and instead take the reported results at face value. Applying such tests typically shrinks the observed estimate. Equally, although the authors did include a risk of bias assessment, the results of that assessment were not employed in any sensitivity analysis (e.g., using bias scores as weights or as a method of exclusion).

End of verification_zhao_2023.html on previous page.

Citing this document

Daniel Nüst (2025). CODECHECK Certificate 2025-005. Zenodo. https://doi.org/10.5281/zenodo.15323262

About CODECHECK

This certificate confirms that the codechecker could independently reproduce the results of a computational analysis given the data and code from a third party. A CODECHECK does not check whether the original computation analysis is correct. However, as all materials required for the reproduction are freely available by following the links in this document, the reader can then study for themselves the code and data.

About this document

This document was created using R Markdown using the `codecheck` R package. `make codecheck.pdf` will regenerate the report file.

```r
sessionInfo()
```

```
## R version 4.5.0 (2025-04-11)
## Platform: x86_64-pc-linux-gnu
## Running under: Ubuntu 22.04.5 LTS
##
## Matrix products: default
## BLAS:   /usr/lib/x86_64-linux-gnu/blas/libblas.so.3.10.0
## LAPACK: /usr/lib/x86_64-linux-gnu/lapack/liblapack.so.3.10.0  LAPACK version 3.10.0
##
## locale:
##  [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C
##  [3] LC_TIME=de_DE.UTF-8        LC_COLLATE=en_US.UTF-8
##  [5] LC_MONETARY=de_DE.UTF-8    LC_MESSAGES=en_US.UTF-8
##  [7] LC_PAPER=de_DE.UTF-8       LC_NAME=C
##  [9] LC_ADDRESS=C               LC_TELEPHONE=C
## [11] LC_MEASUREMENT=de_DE.UTF-8 LC_IDENTIFICATION=C
##
## time zone: Europe/Berlin
## tzcode source: system (glibc)
##
## attached base packages:
## [1] stats     graphics  grDevices datasets  utils
## [6] methods   base
##
## other attached packages:
##  [1] readr_2.1.5      tibble_3.2.1     xtable_1.8-4
##  [4] yaml_2.3.10      rprojroot_2.0.4  knitr_1.50
##  [7] codecheck_0.14.0 parsedate_1.3.2  R.cache_0.16.0
## [10] gh_1.4.1
##
## loaded via a namespace (and not attached):
##  [1] xfun_0.52         rdflib_0.2.9      bspm_0.5.7
##  [4] tzdb_0.5.0        vctrs_0.6.5       tools_4.5.0
##  [7] generics_0.1.3    parallel_4.5.0    curl_6.2.2
## [10] pkgconfig_2.0.3   pdftools_3.5.0    R.oo_1.27.0
## [13] redland_1.0.17-18 assertthat_0.2.1  lifecycle_1.0.4
## [16] git2r_0.36.2      compiler_4.5.0    atom4R_0.3-3
```

```
## [19] stringr_1.5.1     keyring_1.3.2     htmltools_0.5.8.1
## [22] crayon_1.5.3      pillar_1.10.2     whisker_0.4.1
## [25] tidyr_1.3.1       R.utils_2.13.0    cachem_1.1.0
## [28] zen4R_0.10        tidyselect_1.2.1  zip_2.3.2
## [31] digest_0.6.37     stringi_1.8.7     dplyr_1.1.4
## [34] purrr_1.0.4       fastmap_1.2.0     cli_3.6.5
## [37] magrittr_2.0.3    utf8_1.2.4        XML_3.99-0.18
## [40] crul_1.5.0        withr_3.0.2       osfr_0.2.9
## [43] bit64_4.6.0-1     roxygen2_7.3.2    rmarkdown_2.29
## [46] httr_1.4.7        bit_4.6.0         qpdf_1.3.5
## [49] askpass_1.2.1     R.methodsS3_1.8.2 hms_1.1.3
## [52] memoise_2.0.1     evaluate_1.0.3    rlang_1.1.6
## [55] Rcpp_1.0.14       glue_1.8.0        httpcode_0.3.0
## [58] xml2_1.3.8        fauxpas_0.5.2     rorcid_0.7.0
## [61] rstudioapi_0.17.1 vroom_1.6.5       jsonlite_2.0.0
## [64] plyr_1.8.9        R6_2.6.1          fs_1.6.6
```