

# Breathing Huberman

Shravan Vasishth

2023-07-17

tl;dr

- I cannot reproduce a key claim in this paper (Figure 3), that cyclic sighing improves positive feelings compared to mindfulness meditation.
- Surprisingly, about half the data consists of subjects who did not even do the exercise they were assigned. It seems the authors just included these cases, even if the key breathing or meditation task was not carried out.
- A key interaction reported in Figure 3 cannot be significant (it even has two stars), because the reported confidence interval crosses zero.

I suspect that the data provided is not the data used in the published analysis.

I think it is likely that this entire analysis is bogus and in my opinion, this paper should be retracted.

## What the original paper is about

The paper is titled

*Brief structured respiration practices enhance mood and reduce physiological arousal*

and evaluates through an internet-based study the effect of breathwork compared to mindfulness meditation on several different self-perceived well-being metrics.

The paper is available from: <https://doi.org/10.1016/j.xcrm.2022.100895>

It is open access and the data are purportedly available (more on this below).

Mindfulness meditation (hereafter, MM) is compared to three other methods:

- cyclic sighing (CS), “which emphasizes prolonged exhalations”
- box breathing (BB), “which is equal duration of inhalations, breath retentions, and exhalations”
- cyclic hyperventilation (CH) “with retention, with longer inhalations and shorter exhalations”

The main statistical claim I wanted to investigate was:

Using a mixed-effects model, we show that breathwork, especially the exhale-focused cyclic sighing, produces greater improvement in mood ( $p < 0.05$ ) and reduction in respiratory rate ( $p < 0.05$ ) compared with mindfulness meditation. Daily 5-min cyclic sighing has promise as an effective stress management exercise.

Also this:

5 min per day of deliberate breathing practice can cause significant shifts in autonomic tone and well-being.

## Methods

114 participants, with the following between-subjects partitioning:

- MM: 24 (in the data below, I see 25, or 23 if we remove the ones who didn't do the task)
- CS: 30 (in the data provided, I see 31, or 29 if we remove the ones who didn't do the task)
- BB: 21 (in the data provided, I see 22, or 21 if we remove the ones who didn't do the tasks)
- CH: 33 (in the data provided, I see 35, or 34 if we remove the ones who didn't do the tasks)

Dependent variables:

- positive affect (positive and negative affect schedule [PANAS], range 10–50),
- negative affect (PANAS, range 10–50),
- state anxiety (State-Trait Anxiety Inventory [STAI], range 20–80) scores on each participant before and after each breathwork protocol daily.

The experiment was conducted over 29 days, so we have repeated measurements from each subject in each condition, but no repeated measures across the four conditions.

## Claim using linear mixed models (see their Fig 3)

We then examined if breathwork was more effective than mindfulness meditation in reducing anxiety and improving mood. To address this, we constructed a linear mixed-effects model with protocol type and “number of days on protocol” as the fixed effect and participants as the random effect predictors.

the breathwork group had a notably higher increase in daily positive affect (Figures 2A and 2D). The breathwork group also had a significant interaction with the number of days on protocol, such that the daily positive affect increase was larger the more days subjects had been on the protocol (Figures 2A and 2D), suggesting an effect of adherence over time on the daily positive affect benefits.

## The data provided

The authors write:

De-identified raw human physiology and survey data have been deposited at Dryad repository (<https://datadryad.org/>) and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. All original code has been deposited at Zenodo and is publicly available as of the date of publication. DOIs are listed in the key resources table. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Some comments:

- they claim that “all original code” has been deposited at Zenodo. However, all they deposit there is the R preprocessing files, the crucial linear mixed modeling code was done in Matlab and is not provided (at least, I could not find it). No data is provided with the files and no README, so it was impossible for me to run the code.
- They mention a lead contact for obtaining any missing information, and the phrase lead contact is hyper-linked, but the link doesn't lead to any name.
- The data are in an Excel file that is deposited in some other repository (dryad), not Zenodo, and is called:

BWPilot\_CombinedData\_20210803\_fordryad\_addvars\_cleaned\_noround2\_v3.xlsx

The file name is a red flag for me and tells me that the authors are amateurs/beginners in data management. This file name definitely lowers my confidence in this entire study. The file name also raises some questions:

what was in versions 1 and 2, and what exactly does cleaned mean here?

## My data analysis

Read in the csv version of the Excel file:

```
dat<-read.csv("BWPilot_CombinedData_20210803_fordryad_addvars_cleaned_noround2_v3.csv",sep=";",na.strings="")
colnames(dat)
```

```
## [1] "ScrSubjectID"
## [2] "Days.from..Round1.Day1"
## [3] "Round.1.Exercise"
## [4] "Pre.PANAS.Positive.Affect"
## [5] "Pre.PANAS.Negative.Affect"
## [6] "Post.PANAS.Positive.Affect"
## [7] "Post.PANAS.Negative.Affect"
## [8] "Pre.STAI.State.Anxiety"
## [9] "Post.STAI.State.Anxiety"
## [10] "BW.Exercise.Minutes..timer."
## [11] "Did.you.complete.BW.exercises."
## [12] "Sleep.Q.Score"
## [13] "Full.STAI.State.Anxiety"
## [14] "Full.STAI.Trait.Anxiety"
## [15] "RHR"
## [16] "HRV"
## [17] "Sleep.Score"
## [18] "Hours.of.Sleep"
## [19] "Sleep.Efficiency"
## [20] "Respiration.Rate"
## [21] "Sleep.T.Score"
## [22] "Overall..how.easy.was.it.to.use.the.instructions.videos.for.doing.the.intervention."
## [23] "Overall..how.easy.was.it.do.do.the.daily.intervention."
## [24] "subjective.experience..grounding"
## [25] "subjective.experience..focusing"
## [26] "subjective.experience..calming"
## [27] "subjective.experience..energizing"
## [28] "subjective.experience..brief.and.simple"
## [29] "subjective.experience..other.positive"
## [30] "subjective.experience..some.challenges"
```

Isolate the first nine columns plus 11th column for the linear mixed models analyses.

```
dat<-dat[,c(1:9,11)]
head(dat)
```

```
## ScrSubjectID Days.from..Round1.Day1 Round.1.Exercise
## 1 1000131 -28 Box Breathing
## 2 1000131 -27 Box Breathing
## 3 1000131 -26 Box Breathing
## 4 1000131 -25 Box Breathing
## 5 1000131 -24 Box Breathing
## 6 1000131 -23 Box Breathing
## Pre.PANAS.Positive.Affect Pre.PANAS.Negative.Affect
## 1 NA NA
## 2 NA NA
## 3 NA NA
```

```
## 4          NA          NA
## 5          NA          NA
## 6          NA          NA
##   Post.PANAS.Positive.Affect Post.PANAS.Negative.Affect Pre.STAI.State.Anxiety
## 1          NA          NA          NA
## 2          NA          NA          NA
## 3          NA          NA          NA
## 4          NA          NA          NA
## 5          NA          NA          NA
## 6          NA          NA          NA
##   Post.STAI.State.Anxiety Did.you.complete.BW.exercises.
## 1          NA          NA
## 2          NA          NA
## 3          NA          NA
## 4          NA          NA
## 5          NA          NA
## 6          NA          NA
```

Rename the columns to more tractable names:

```
colnames(dat)<-c("subj","days","exercise","PrePANASPos","PrePANASNeg","PostPANASPos","PostPANASNeg","PreSTAI","PostSTAI","completed")
head(dat)
```

```
##      subj days      exercise PrePANASPos PrePANASNeg PostPANASPos PostPANASNeg
## 1 1000131 -28 Box Breathing          NA          NA          NA          NA
## 2 1000131 -27 Box Breathing          NA          NA          NA          NA
## 3 1000131 -26 Box Breathing          NA          NA          NA          NA
## 4 1000131 -25 Box Breathing          NA          NA          NA          NA
## 5 1000131 -24 Box Breathing          NA          NA          NA          NA
## 6 1000131 -23 Box Breathing          NA          NA          NA          NA
##   PreSTAI PostSTAI completed
## 1      NA      NA      NA
## 2      NA      NA      NA
## 3      NA      NA      NA
## 4      NA      NA      NA
## 5      NA      NA      NA
## 6      NA      NA      NA
```

Remove all irrelevant days:

```
dat<-subset(dat,days>0 & days < 29)
head(dat)
```

```
##      subj days      exercise PrePANASPos PrePANASNeg PostPANASPos PostPANASNeg
## 30 1000131   1 Box Breathing          32          24          25          13
## 31 1000131   2 Box Breathing          24          14          26          12
## 32 1000131   3 Box Breathing          30          14          30          10
## 33 1000131   4 Box Breathing          23          10          29          10
## 34 1000131   5 Box Breathing          18          14          16          13
## 35 1000131   6 Box Breathing          18          12          22          10
##   PreSTAI PostSTAI completed
## 30     48     33      NA
## 31     42     26      NA
## 32     33     25      NA
## 33     27     28      NA
## 34     41     31      NA
## 35     37     27       1
```

Sanity check:

```
length(unique(dat$subj))
```

```
## [1] 113
```

One subject seems to be missing: the authors wrote they had 114 subjects.

Convert the treatments to factors:

```
dat$exercise<-factor(dat$exercise)
```

The authors write that the names for the exercises in the data file do not correspond to the names used in the paper:

Round 1 Exercise: Participant's assigned intervention. Note the nomenclature is slightly different than the manuscript. Super Oxygenation = Cyclic Hyperventilation with Retention. Slow Breathing = Cyclic Sighing

We will therefore rename the conditions from those in the data file to those in the paper:

```
dat$exercise<-factor(dat$exercise,
                     levels=c("Mindful Meditation",
                              "Box Breathing",
                              "Slow Breathing",
                              "Super Oxygenation"))

dat$cond<-ifelse(dat$exercise=="Mindful Meditation",
                 "MM",
                 ifelse(dat$exercise=="Box Breathing",
                       "BB",
                       ifelse(dat$exercise=="Slow Breathing",
                             "CS",
                             ifelse(dat$exercise=="Super Oxygenation",
                                     "CH",NA))))
```

Next, set the factor levels so that the baseline condition is MM (Mindfulness Meditation), because this is compared to the other three conditions:

```
dat$cond<-factor(dat$cond,levels=c("MM","CS","BB","CH"))
contrasts(dat$cond)
```

```
##      CS BB CH
## MM   0  0  0
## CS   1  0  0
## BB   0  1  0
## CH   0  0  1
```

The number of subjects in each condition:

```
length(unique(subset(dat,cond=="MM")$subj))
```

```
## [1] 25
```

```
length(unique(subset(dat,cond=="CS")$subj))
```

```
## [1] 31
```

```
length(unique(subset(dat,cond=="BB")$subj))
```

```
## [1] 22
```

```
length(unique(subset(dat,cond=="CH")$subj))
```

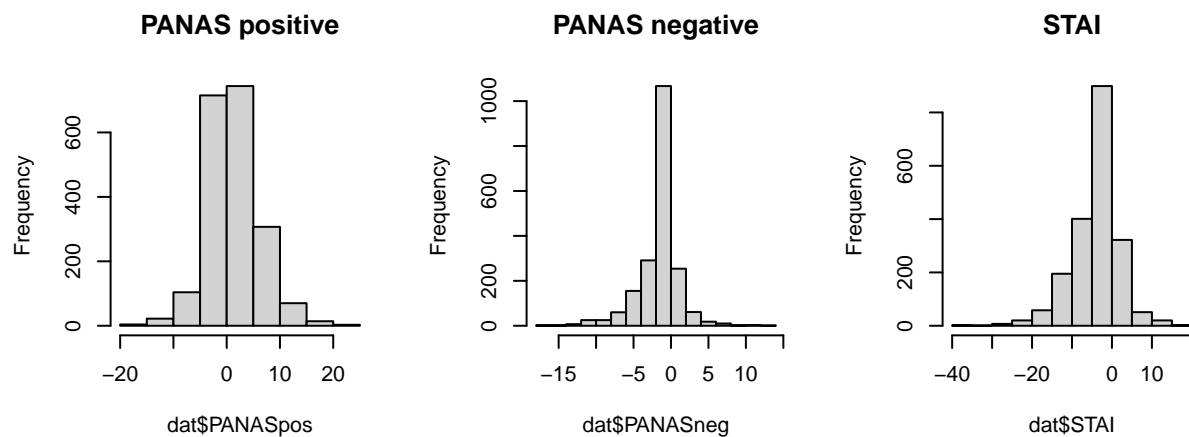
```
## [1] 35
```

Next, set up the dependent variables for Figures 2 and 3:

```
## create dependent variables:
dat$PANASpos<-dat$PostPANASpos-dat$PrePANASpos
dat$PANASneg<-dat$PostPANASneg-dat$PrePANASneg
dat$STAI <- dat$PostSTAI-dat$PreSTAI
```

Graphical visualization of the dependent variables:

```
op<-par(mfrow=c(1,3),pty="s")
hist(dat$PANASpos,main="PANAS positive")
hist(dat$PANASneg,main="PANAS negative")
hist(dat$STAI,main="STAI")
```



Next, check the structure of the data:

The key manipulation is between subjects:

```
xtabs(~subj+cond,dat)
```

```
##      cond
## subj  MM CS BB CH
## 1000131 0 0 28 0
## 1000262 0 0 28 0
## 1000393 0 28 0 0
## 1000524 0 0 0 28
## 1000655 0 0 28 0
## 1001048 28 0 0 0
## 1001310 0 0 28 0
## 1001441 28 0 0 0
## 1001572 0 28 0 0
## 1001703 28 0 0 0
## 1001834 0 28 0 0
## 1002751 0 28 0 0
## 1002882 0 0 0 28
## 1003013 0 28 0 0
## 1003144 28 0 0 0
## 1004192 0 0 28 0
## 1004323 0 28 0 0
```

##	1004454	0	0	0	27
##	1004585	0	0	0	28
##	1004978	28	0	0	0
##	1005240	0	0	28	0
##	1005764	0	28	0	0
##	1005895	0	0	0	28
##	1006026	0	0	28	0
##	1006157	0	28	0	0
##	1006550	0	0	28	0
##	1007205	0	28	0	0
##	1007336	28	0	0	0
##	1007598	0	0	0	28
##	1008122	0	0	0	28
##	1008253	0	0	0	28
##	1008384	0	28	0	0
##	1008515	28	0	0	0
##	1008646	0	27	0	0
##	1008777	0	0	0	27
##	1008908	28	0	0	0
##	1009039	0	0	0	28
##	1009170	28	0	0	0
##	1009432	0	28	0	0
##	1009956	0	28	0	0
##	1010087	0	28	0	0
##	1010218	0	28	0	0
##	1010611	0	0	28	0
##	1011004	28	0	0	0
##	1011659	0	28	0	0
##	1011790	28	0	0	0
##	1012052	0	0	28	0
##	1012576	0	28	0	0
##	1012707	0	0	0	28
##	1013362	28	0	0	0
##	1013493	28	0	0	0
##	1014672	28	0	0	0
##	1015065	0	0	0	28
##	1015196	28	0	0	0
##	1015589	0	28	0	0
##	1015851	0	0	0	28
##	1015982	0	0	0	28
##	1016244	0	0	28	0
##	1016375	0	0	0	28
##	1016637	28	0	0	0
##	1016768	0	27	0	0
##	1016899	0	28	0	0
##	1017030	0	28	0	0
##	1017292	0	0	0	28
##	1017423	0	0	0	28
##	1017554	28	0	0	0
##	1018209	0	0	28	0
##	1018471	0	0	28	0
##	1018995	0	0	27	0
##	1019126	0	28	0	0
##	1019519	0	0	0	28

```
## 1019912 0 0 28 0
## 1020043 0 0 28 0
## 1020436 0 0 0 27
## 1020698 0 28 0 0
## 1020960 0 0 28 0
## 1021222 0 0 0 23
## 1021353 0 28 0 0
## 1021615 0 0 28 0
## 1021746 0 0 0 28
## 1022270 27 0 0 0
## 1022401 0 0 0 28
## 1022532 0 0 0 28
## 1023056 28 0 0 0
## 1023318 28 0 0 0
## 1023580 0 0 28 0
## 1023711 0 28 0 0
## 1023973 0 28 0 0
## 1024104 0 0 0 27
## 1024497 0 0 0 28
## 1024890 0 0 0 28
## 1025545 28 0 0 0
## 1025807 0 28 0 0
## 1025938 0 28 0 0
## 1026331 0 0 0 28
## 1026593 0 0 0 28
## 1027117 0 0 0 28
## 1027248 28 0 0 0
## 1027510 27 0 0 0
## 1028034 0 28 0 0
## 1028296 0 0 28 0
## 1028689 0 0 0 27
## 1028951 0 0 0 28
## 1029082 0 0 0 28
## 1029606 28 0 0 0
## 1029868 0 0 28 0
## 1030130 28 0 0 0
## 1030261 0 27 0 0
## 1030654 0 0 0 28
## 1030785 0 0 28 0
## 1031047 0 0 0 28
## 1031178 0 0 0 28
## 1031309 0 28 0 0
```

We have at most one data point per day (some missing data):

```
xtabs(~subj+days,dat)
```

```
##          days
## subj      1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
## 1000131 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
## 1000262 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
## 1000393 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
## 1000524 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
## 1000655 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
## 1001048 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
```



[illegible]



##	subj	27	28
##	1000131	1	1
##	1000262	1	1
##	1000393	1	1
##	1000524	1	1
##	1000655	1	1
##	1001048	1	1
##	1001310	1	1
##	1001441	1	1
##	1001572	1	1
##	1001703	1	1
##	1001834	1	1
##	1002751	1	1
##	1002882	1	1
##	1003013	1	1
##	1003144	1	1
##	1004192	1	1
##	1004323	1	1
##	1004454	1	0
##	1004585	1	1
##	1004978	1	1
##	1005240	1	1
##	1005764	1	1
##	1005895	1	1
##	1006026	1	1
##	1006157	1	1
##	1006550	1	1
##	1007205	1	1
##	1007336	1	1
##	1007598	1	1
##	1008122	1	1
##	1008253	1	1
##	1008384	1	1
##	1008515	1	1
##	1008646	1	0
##	1008777	1	0
##	1008908	1	1
##	1009039	1	1
##	1009170	1	1
##	1009432	1	1
##	1009956	1	1
##	1010087	1	1
##	1010218	1	1
##	1010611	1	1
##	1011004	1	1
##	1011659	1	1
##	1011790	1	1
##	1012052	1	1
##	1012576	1	1
##	1012707	1	1
##	1013362	1	1
##	1013493	1	1
##	1014672	1	1
##	1015065	1	1

##	1015196	1	1
##	1015589	1	1
##	1015851	1	1
##	1015982	1	1
##	1016244	1	1
##	1016375	1	1
##	1016637	1	1
##	1016768	1	0
##	1016899	1	1
##	1017030	1	1
##	1017292	1	1
##	1017423	1	1
##	1017554	1	1
##	1018209	1	1
##	1018471	1	1
##	1018995	1	0
##	1019126	1	1
##	1019519	1	1
##	1019912	1	1
##	1020043	1	1
##	1020436	1	0
##	1020698	1	1
##	1020960	1	1
##	1021222	1	0
##	1021353	1	1
##	1021615	1	1
##	1021746	1	1
##	1022270	1	0
##	1022401	1	1
##	1022532	1	1
##	1023056	1	1
##	1023318	1	1
##	1023580	1	1
##	1023711	1	1
##	1023973	1	1
##	1024104	1	0
##	1024497	1	1
##	1024890	1	1
##	1025545	1	1
##	1025807	1	1
##	1025938	1	1
##	1026331	1	1
##	1026593	1	1
##	1027117	1	1
##	1027248	1	1
##	1027510	1	0
##	1028034	1	1
##	1028296	1	1
##	1028689	1	0
##	1028951	1	1
##	1029082	1	1
##	1029606	1	1
##	1029868	1	1
##	1030130	1	1

```
## 1030261 1 0
## 1030654 1 1
## 1030785 1 1
## 1031047 1 1
## 1031178 1 1
## 1031309 1 1
```

## Subjects did the task only in about 50% of the cases

The paper states:

During the 28-day intervention period, participants did their assigned 5-min exercise and completed two questionnaires before and after, the State Anxiety Inventory<sup>60</sup> and the Positive and Negative Affect Schedule (PANAS).

The paper also states:

Participants were assumed to have completed the breathing protocol if they had filled out the pre- and post-measures for a particular day.

But it seems some subjects did not complete the task; some didn't do the task at all!:

```
xtabs(~subj+completed,dat)
```

```
##           completed
## subj      0      1
## 1000131    1    22
## 1000262   10   14
## 1000393    6   17
## 1000524    9   18
## 1000655   15   11
## 1001048   20    7
## 1001310    6   18
## 1001441    0   23
## 1001572    1   22
## 1001703    9   19
## 1001834    3   20
## 1002751   23    3
## 1002882    9   14
## 1003013   23    0
## 1003144    2   21
## 1004192   12   11
## 1004323   13   10
## 1004454   12   12
## 1004585   12   16
## 1004978   12   11
## 1005240   10   18
## 1005764   11   15
## 1005895    6   17
## 1006026    3   20
## 1006157   11   12
## 1006550   10   13
## 1007205   23    5
## 1007336    3   20
## 1007598    9   17
## 1008122   11   15
## 1008253    0   23
```

##	1008384	2	21
##	1008515	4	19
##	1008646	26	1
##	1008777	15	9
##	1008908	7	21
##	1009039	21	2
##	1009170	23	3
##	1009432	13	13
##	1009956	8	15
##	1010087	7	16
##	1010218	10	13
##	1010611	6	18
##	1011004	10	13
##	1011659	5	18
##	1011790	25	1
##	1012052	0	23
##	1012576	22	1
##	1012707	6	17
##	1013362	10	18
##	1013493	26	0
##	1014672	12	14
##	1015065	9	17
##	1015196	4	19
##	1015589	2	21
##	1015851	19	8
##	1015982	10	18
##	1016244	13	10
##	1016375	2	21
##	1016637	0	23
##	1016768	12	15
##	1016899	4	18
##	1017030	5	23
##	1017292	3	24
##	1017423	1	27
##	1017554	28	0
##	1018209	1	22
##	1018471	2	21
##	1018995	21	6
##	1019126	14	14
##	1019519	11	13
##	1019912	25	3
##	1020043	23	0
##	1020436	22	1
##	1020698	14	13
##	1020960	6	22
##	1021222	17	2
##	1021353	2	24
##	1021615	3	20
##	1021746	16	7
##	1022270	10	13
##	1022401	4	18
##	1022532	0	27
##	1023056	3	24
##	1023318	17	10

```
## 1023580 4 24
## 1023711 27 0
## 1023973 2 25
## 1024104 27 0
## 1024497 1 19
## 1024890 1 26
## 1025545 5 18
## 1025807 7 21
## 1025938 4 24
## 1026331 3 19
## 1026593 21 7
## 1027117 2 25
## 1027248 25 2
## 1027510 15 9
## 1028034 6 17
## 1028296 8 16
## 1028689 20 7
## 1028951 9 18
## 1029082 18 8
## 1029606 3 25
## 1029868 3 24
## 1030130 14 12
## 1030261 26 1
## 1030654 2 25
## 1030785 20 8
## 1031047 2 26
## 1031178 4 24
## 1031309 8 20
```

If we remove the data where the task was not completed, we would lose about half the data. However, if subjects did not do the task assigned, should they be included in the analysis? I would say no, but I will report linear mixed models with and without these non-compliant subjects, because it seems that the authors ignored the fact that many subjects didn't even do the assigned task (they ignored the completed column).

```
dim(dat)
```

```
## [1] 3148 14
```

```
dat_completed<-subset(dat,completed==1)
dim(dat_completed)
```

```
## [1] 1674 14
```

The number of subjects in each condition after we remove subjects who didn't do the task:

```
length(unique(subset(dat_completed,cond=="MM")$subj))
```

```
## [1] 23
```

```
length(unique(subset(dat_completed,cond=="CS")$subj))
```

```
## [1] 29
```

```
length(unique(subset(dat_completed,cond=="BB")$subj))
```

```
## [1] 21
```

```
length(unique(subset(dat_completed,cond=="CH")$subj))
```

## [1] 34

### Figure 3: approximate recreations

Here is their Figure 3:

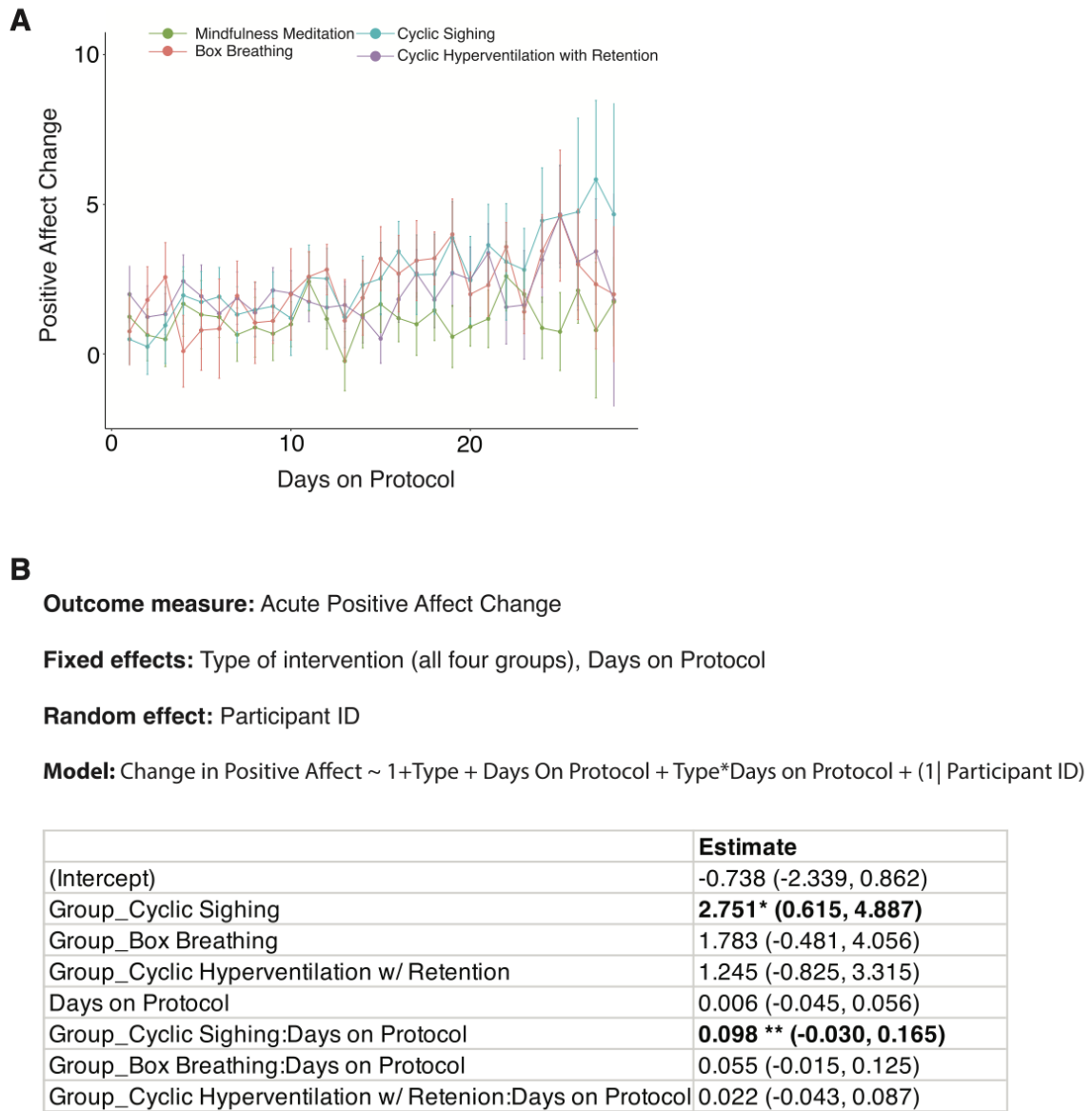


Figure 1: figure 3

I can reproduce their Figure 3 more or less, at least it looks visually similar to their plot. In particular, the CS (cyclic sighing) values seem to rise higher than the mindfulness meditation (MM) values over time.

This plot is based on the full data, regardless of whether subjects completed the task:

```
op<-par(mfrow=c(2,2),pty="s")
```



```

conditions<-levels(dat$cond)

# for each condition, make a plot:
for(j in 1:length(conditions)){
  # First, aggregate the data:
  dat_agg<-with(subset(dat,cond==conditions[j]),
    tapply(PANASpos,days,mean,na.rm=TRUE))

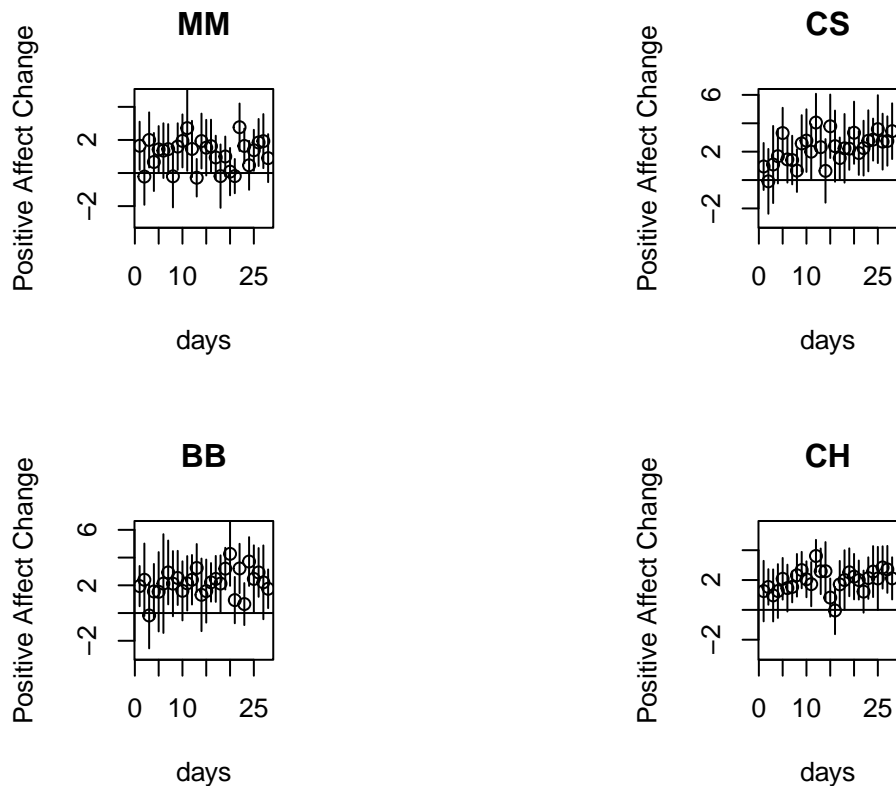
  # Compute SD:
  dat_agg_SD<-with(subset(dat,cond==conditions[j]),
    tapply(PANASpos,days,sd,na.rm=TRUE))
  # figure out no. of data points in the condition being plotted:
  lengths<-rep(NA,28)
  for(i in 1:28){
    lengths[i]<-dim(subset(dat,cond==conditions[j] & days == i))[1]
  }

  # compute SE for plotting confidence intervals:
  SE<-dat_agg_SD/sqrt(lengths)

  datframe<-data.frame(y=dat_agg,
    x=1:28,
    lower=dat_agg-2*SE,
    upper=dat_agg+2*SE)

  plot(datframe$x,datframe$y,ylim=c(-3,max(datframe$y)+2),
    xlab="days",
    ylab="Positive Affect Change",
    main=conditions[j])
  with(datframe,
    arrows(x0=x,y0=lower,x1=x,y1=upper,angle=90,
      length=0)
  )
  abline(h=0)
}

```



Here is the plot using only those data in which subjects actually did the task:

```
op<-par(mfrow=c(2,2),pty="s")

conditions<-levels(dat_completed$cond)

# for each condition, make a plot:
for(j in 1:length(conditions)){
  # First, aggregate the data:
  dat_agg<-with(subset(dat_completed,cond==conditions[j]),
    tapply(PANASpos,days,mean,na.rm=TRUE))

  # Compute SD:
  dat_agg_SD<-with(subset(dat_completed,cond==conditions[j]),
    tapply(PANASpos,days,sd,na.rm=TRUE))

  # figure out no. of data points in the condition being plotted:
  lengths<-rep(NA,28)
  for(i in 1:28){
    lengths[i]<-dim(subset(dat_completed,cond==conditions[j] & days == i))[1]
  }

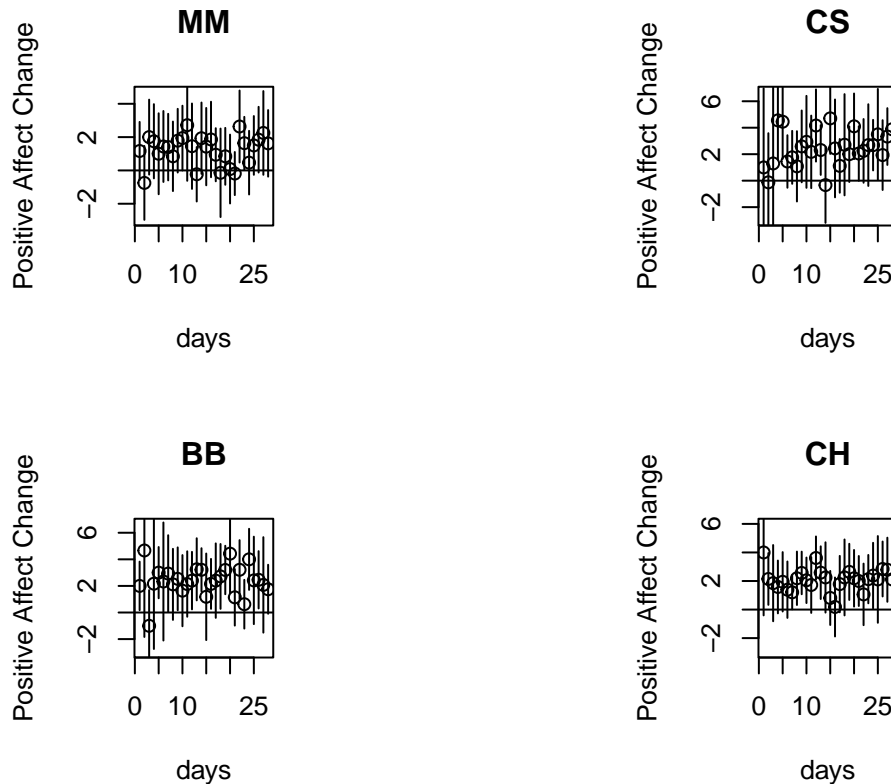
  # compute SE for plotting confidence intervals:
  SE<-dat_agg_SD/sqrt(lengths)

  dataframe<-data.frame(y=dat_agg,
    x=1:28,
    lower=dat_agg-2*SE,
    upper=dat_agg+2*SE)
```

```

plot(datframe$x,datframe$y,ylim=c(-3,max(datframe$y)+2),
     xlab="days",
     ylab="Positive Affect Change",
     main=conditions[j])
with(datframe,
arrows(x0=x,y0=lower,x1=x,y1=upper,angle=90,
      length=0)
)
abline(h=0)
}

```



**Figure 3** linear mixed model with all data (regardless of whether the subject did the task):

The paper states

“A mixed-effects modeling approach was used to compare changes across groups (Figure 2). Daily change between pre and post protocol for each subject was used as the main unit for modeling. All variables were centered by subtracting the mean before feeding into the model. The cumulative day variable was centered at day 28. Data processing was performed in R and linear mixed-effects modeling was conducted using the “fitlme” function in MATLAB.”

I don't have Matlab so I used lme4.

```
library(lme4)
```

```
## Loading required package: Matrix
```

```
# center days around 28:
```

```
dat$days<-dat$days-28
```

```

m<-lmer(PANASpos~cond*days + (1|subj),data = dat)
summary(m)

## Linear mixed model fit by REML ['lmerMod']
## Formula: PANASpos ~ cond * days + (1 | subj)
## Data: dat
##
## REML criterion at convergence: 10953.3
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -5.1268 -0.4686  0.0269  0.5407  4.0591
##
## Random effects:
## Groups   Name                Variance Std.Dev.
## subj     (Intercept)    9.699     3.114
## Residual                    12.666     3.559
## Number of obs: 1983, groups:  subj, 108
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  1.093524   0.759559   1.440
## condCS       1.910585   1.008148   1.895
## condBB       1.303898   1.084911   1.202
## condCH       1.346777   0.980427   1.374
## days        -0.003793   0.022570  -0.168
## condCS:days  0.070344   0.029555   2.380
## condBB:days  0.034093   0.031345   1.088
## condCH:days  0.035500   0.028849   1.231
##
## Correlation of Fixed Effects:
##              (Intr) condCS condBB condCH days   cndCS: cndBB:
## condCS      -0.753
## condBB      -0.700  0.527
## condCH      -0.775  0.584  0.542
## days         0.428 -0.322 -0.300 -0.332
## condCS:days -0.327  0.425  0.229  0.253 -0.764
## condBB:days -0.308  0.232  0.415  0.239 -0.720  0.550
## condCH:days -0.335  0.252  0.234  0.429 -0.782  0.597  0.563

```

**Figure 3** linear mixed model with only the data from subjects who completed the task:

```

dat_completed$days<-dat_completed$days-28

m_completed<-lmer(PANASpos~cond*days + (1|subj),data = dat_completed)
summary(m_completed)

## Linear mixed model fit by REML ['lmerMod']
## Formula: PANASpos ~ cond * days + (1 | subj)
## Data: dat_completed
##
## REML criterion at convergence: 9058.4

```

```
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -5.2310 -0.4652  0.0047  0.5085  4.2177
##
## Random effects:
##   Groups   Name      Variance Std.Dev.
##   subj     (Intercept) 10.13    3.183
##   Residual             11.62    3.409
## Number of obs: 1659, groups:  subj, 107
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  0.98858    0.78513   1.259
## condCS       1.90587    1.05005   1.815
## condBB       1.19978    1.11909   1.072
## condCH       1.28346    1.00997   1.271
## days        -0.01751    0.02620  -0.669
## condCS:days  0.06970    0.03451   2.020
## condBB:days  0.02322    0.03692   0.629
## condCH:days  0.02209    0.03281   0.673
##
## Correlation of Fixed Effects:
##              (Intr) condCS condBB condCH days   cndCS: cndBB:
## condCS       -0.748
## condBB       -0.702  0.525
## condCH       -0.777  0.581  0.545
## days         0.441 -0.330 -0.309 -0.343
## condCS:days -0.335  0.433  0.235  0.260 -0.759
## condBB:days -0.313  0.234  0.428  0.243 -0.710  0.539
## condCH:days -0.352  0.263  0.247  0.433 -0.798  0.606  0.567
```

## Discussion of linear mixed model results

Discrepancies between my analysis and the one in the paper:

- Discrepancy 1: In the paper, figure 3 shows a significant effect of CS vs MM: the claimed estimate is 2.751 [0.651, 4.887]. My estimate is not even close (see condCS in the model outputs).
- Discrepancy 2: In the paper, a significant (!) interaction is reported between days and CSvsMM: 0.098 [-0.030, 0.0165]. This interaction is even more significant than the first one above, as there are two significance stars attached to it. But this interaction **cannot** be significant in the paper, because the confidence interval crosses 0. However, my interaction term **is** significant, but with different numbers (see condCS:days).

## Design analysis a la Gelman and Carlin

It's useful to try to figure out what sample size one might need for the above comparisons. I focus only on MM vs CS.

```
## design analysis
## sample sizes in each group:
nMM<-length(unique(subset(dat,cond=="MM")$subj))
nCS<-length(unique(subset(dat,cond=="CS")$subj))

## effect size as reported (not reproducible):
```

```
d<-2.751
## confidence intervals:
upper<-4.887
lower<-0.615
## figure out SE:
SE<-(upper-lower)/4
```

Prospective power for the reported effect of CS vs MM in Figure 3's linear mixed model analysis, using the `power.t.test` function:

We first need to figure out the standard deviation:

```
## assume identical SEs in each group and compute std dev from it:
stddev<-sqrt(SE^2*nMM + SE^2*nCS)
```

What sample size would we need to achieve the reported effect size with power 0.80?

```
## carry out power test using reported mean effect:
power.t.test(d=d,sd=stddev,type="two.sample",alternative="two.sided",
             power=0.80,
             strict=TRUE)
```

```
##
##      Two-sample t test power calculation
##
##              n = 133.4581
##              delta = 2.751
##              sd = 7.99218
##              sig.level = 0.05
##              power = 0.8
##              alternative = two.sided
##
## NOTE: n is number in *each* group
```

```
## carry out power test using lower bound of reported mean effect:
power.t.test(d=lower,sd=stddev,type="two.sample",alternative="two.sided",
             power=0.80,
             strict=TRUE)
```

```
##
##      Two-sample t test power calculation
##
##              n = 2652.005
##              delta = 0.615
##              sd = 7.99218
##              sig.level = 0.05
##              power = 0.8
##              alternative = two.sided
##
## NOTE: n is number in *each* group
```

```
## carry out power test using upper bound of reported mean effect:
power.t.test(d=upper,sd=stddev,type="two.sample",alternative="two.sided",
             power=0.80,
             strict=TRUE)
```

```
##
```

```
##      Two-sample t test power calculation
##
##          n = 42.96567
##        delta = 4.887
##          sd = 7.99218
##    sig.level = 0.05
##        power = 0.8
##    alternative = two.sided
##
## NOTE: n is number in each group
```

Conclusion: the study reported would need about 135 subjects in each arm, probably even more (2652 in each arm) if the effect size is as small as the lower bound of the 95% confidence interval reported in the paper.