

My journey from error messages to R packages

Building tools to increase reproducibility
and productivity in clinical psychology

Milan Wiedemann



milanwiedemann



@milanwiedemann

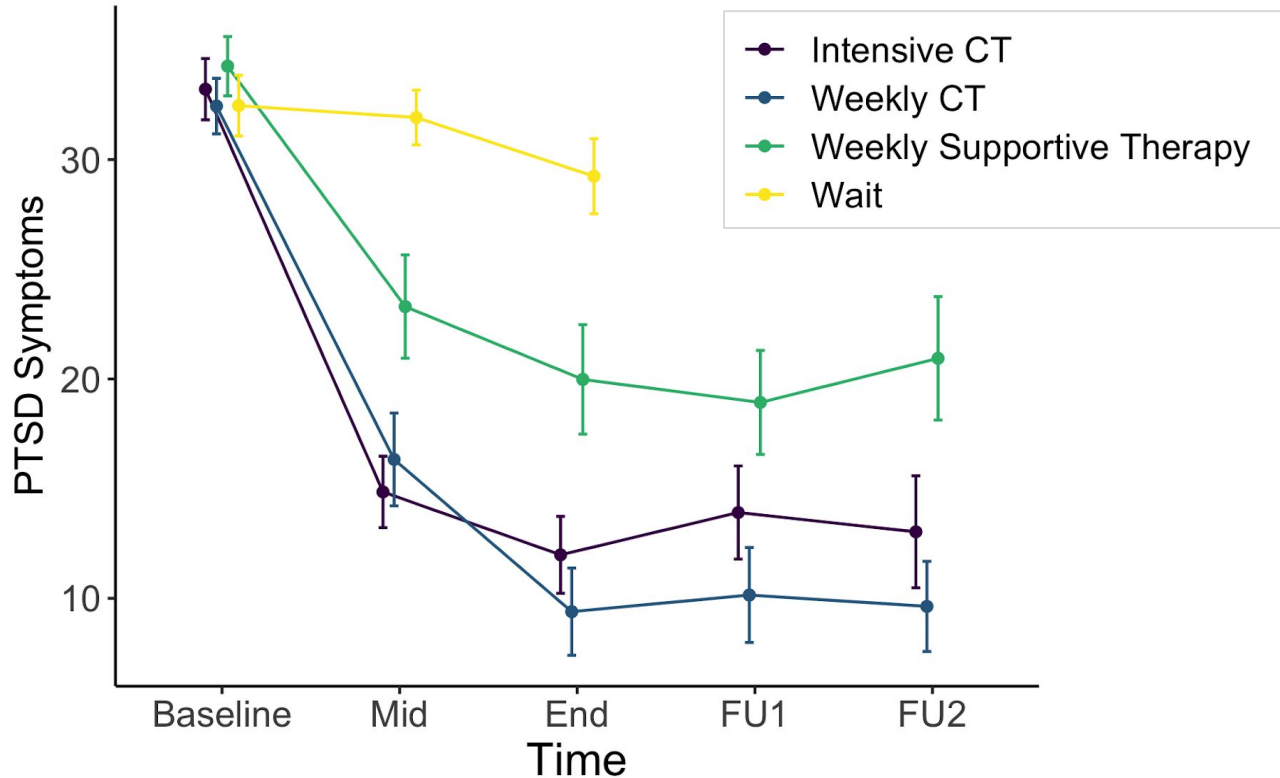
R Oxford User Group

2 Feb 2020

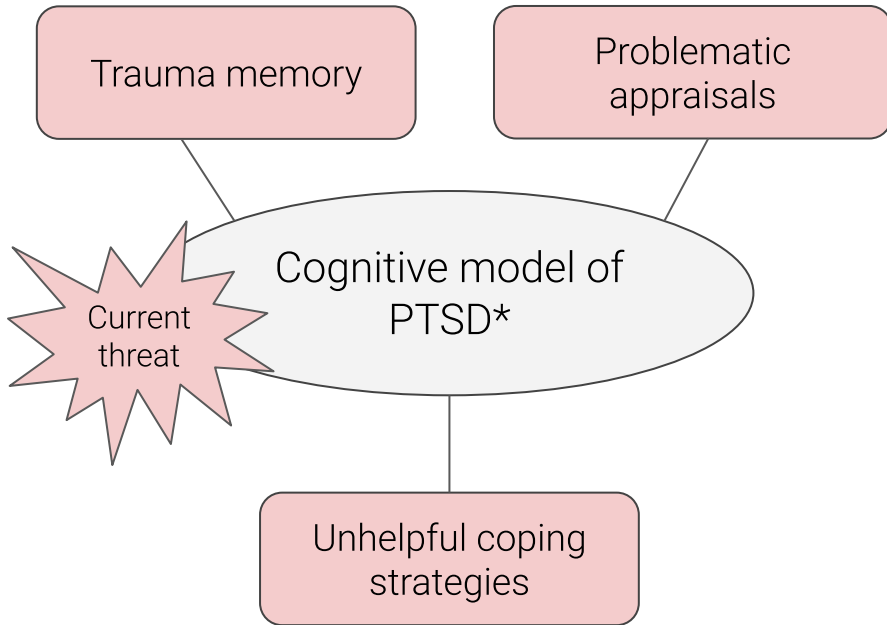
Oxford Centre for Anxiety Disorders and Trauma



Yes! But what's the mechanism?

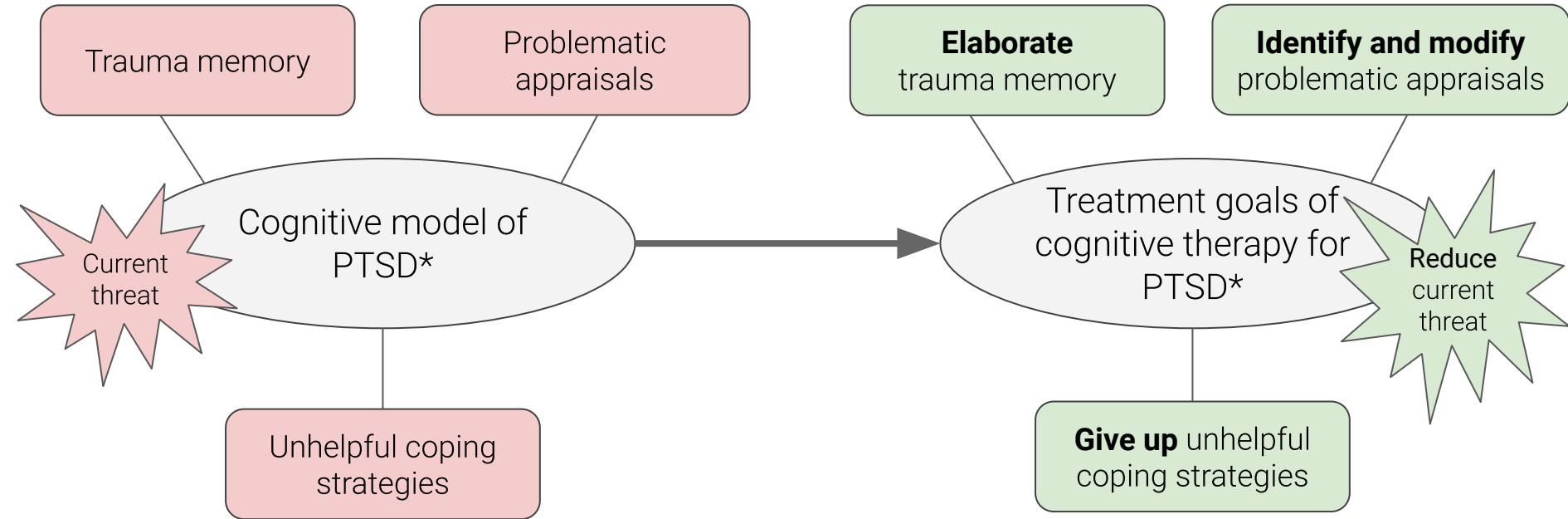


From theory to therapy and back



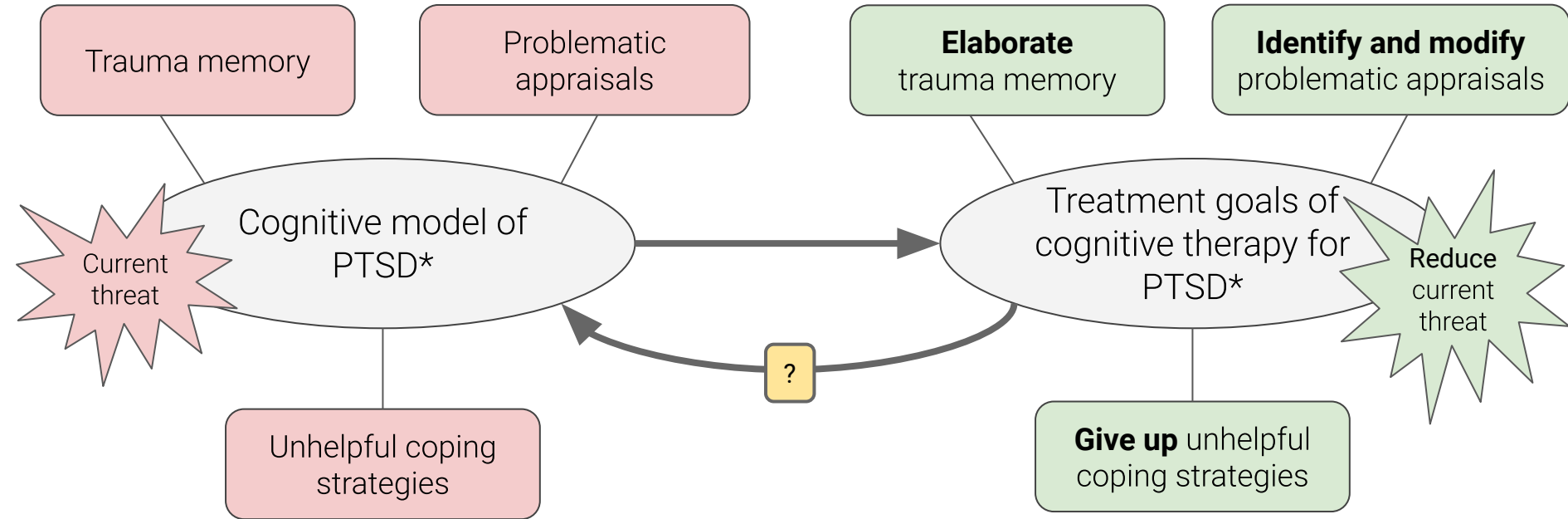
* Based on Ehlers and Clark (2000) and Ehlers et al. (2005).

From theory to therapy and back



* Based on Ehlers and Clark (2000) and Ehlers et al. (2005).

From theory to therapy and back



* Based on Ehlers and Clark (2000) and Ehlers et al. (2005).

Milestones during my R journey



CRAN



```
my_fun <- function(...) {  
  trying stuff until it  
  works  
}
```



blabla ... R Package ...



Project 2: Modeling Change

Project 1: Sudden Gains

Year 1

Year 2

Year 3

Year 4

Problem

Sudden Gains

Our Solution

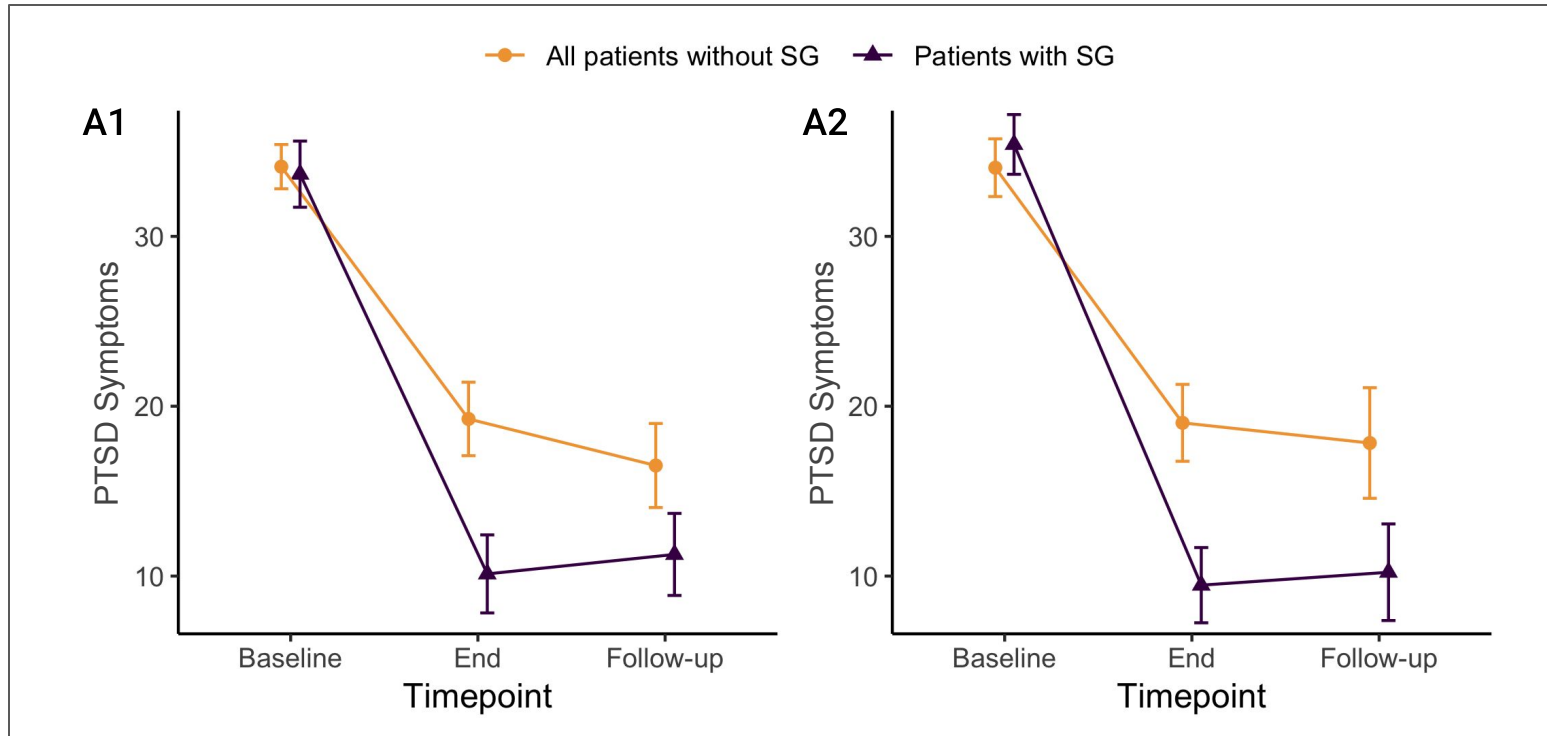
Problem

Modeling Change

Our Solution

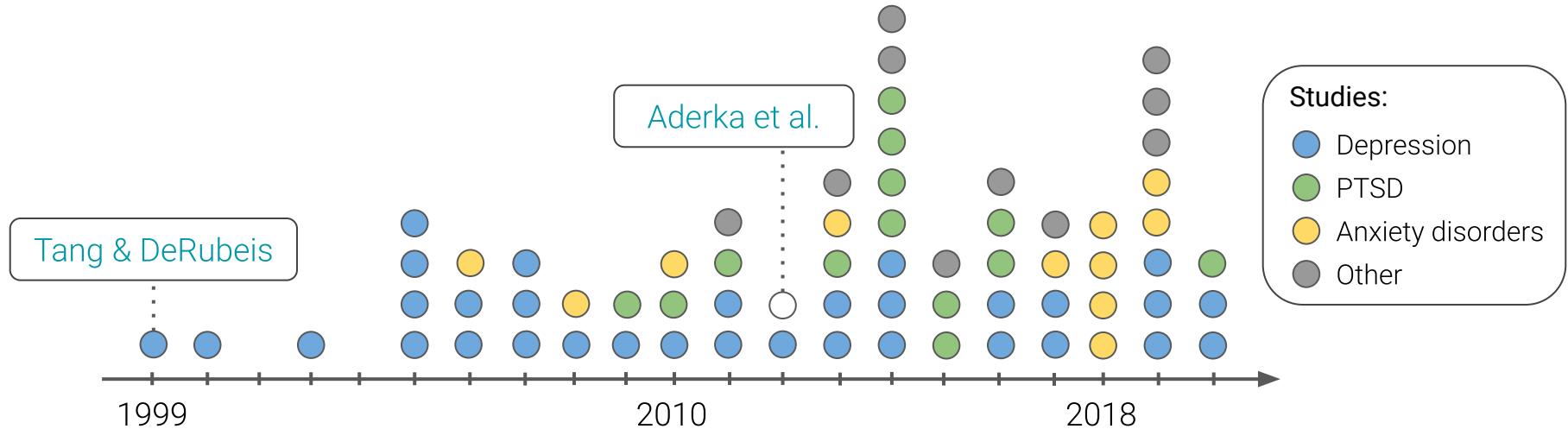
Sudden Gains

Why are sudden gains interesting?



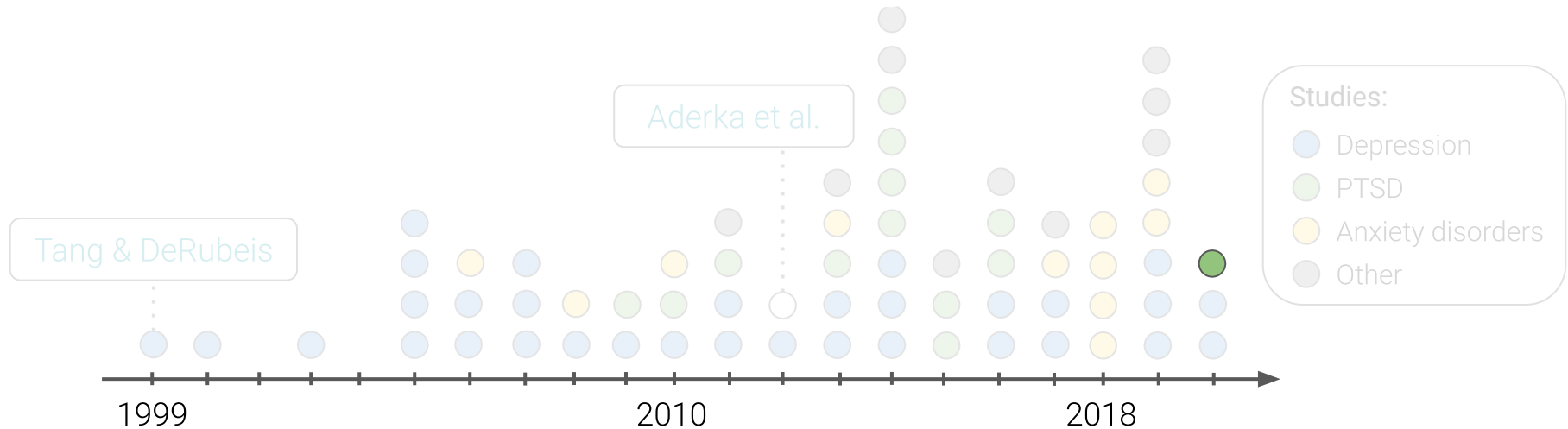
* Results from Wiedemann et al. (2020). SG = Sudden Gain. Sample 1 (A1): $n = 248$, Sample 2 (A2): $n = 234$.

Sudden gains studies in the literature



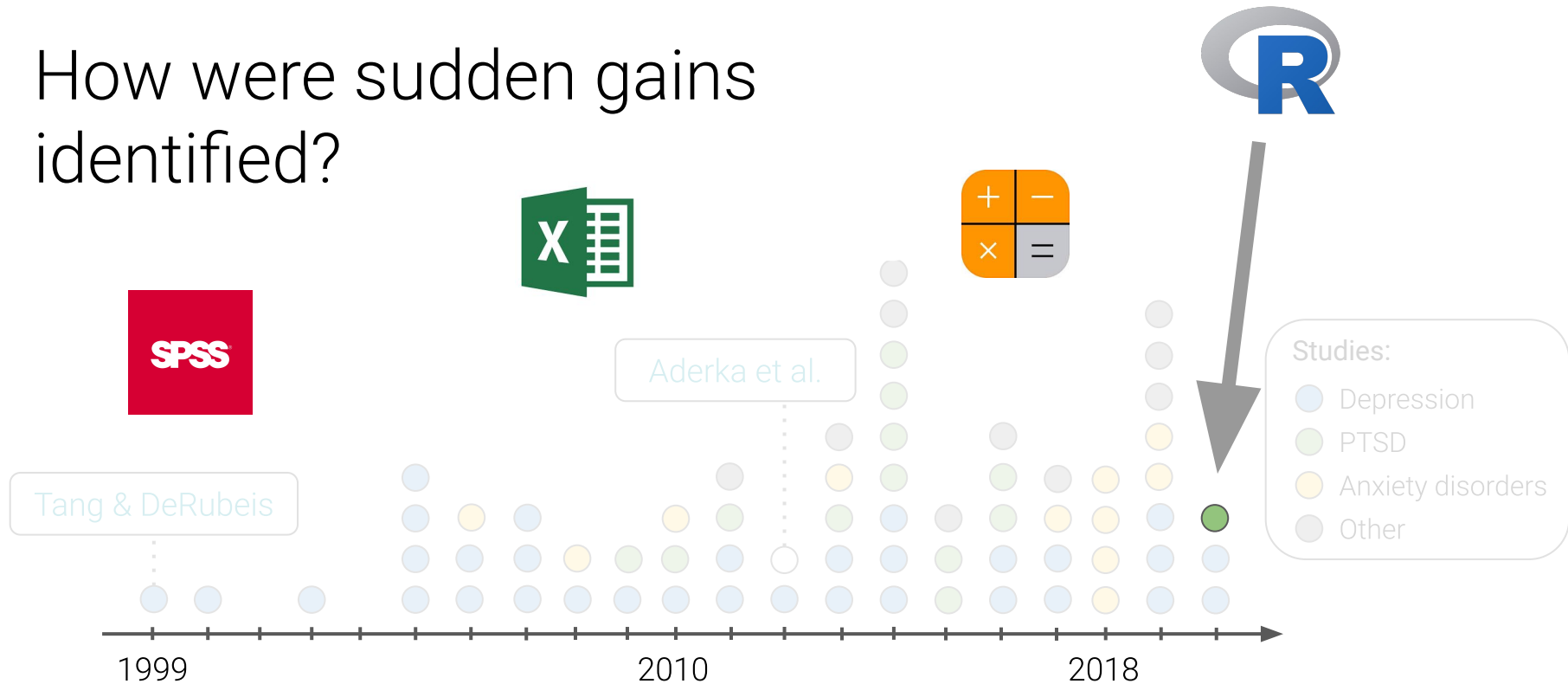
Sudden gains studies with open code

How were sudden gains identified?

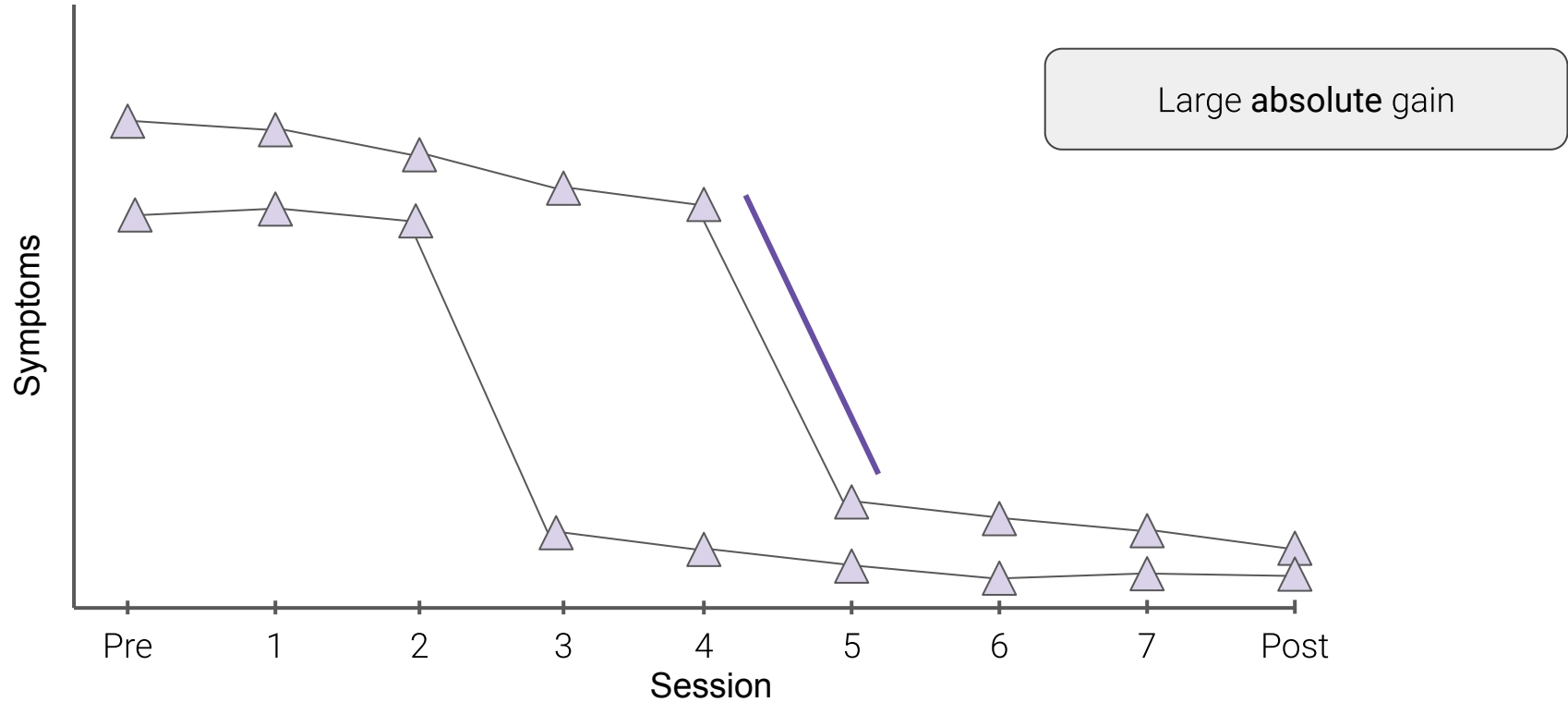


Sudden gains studies with open code

How were sudden gains identified?

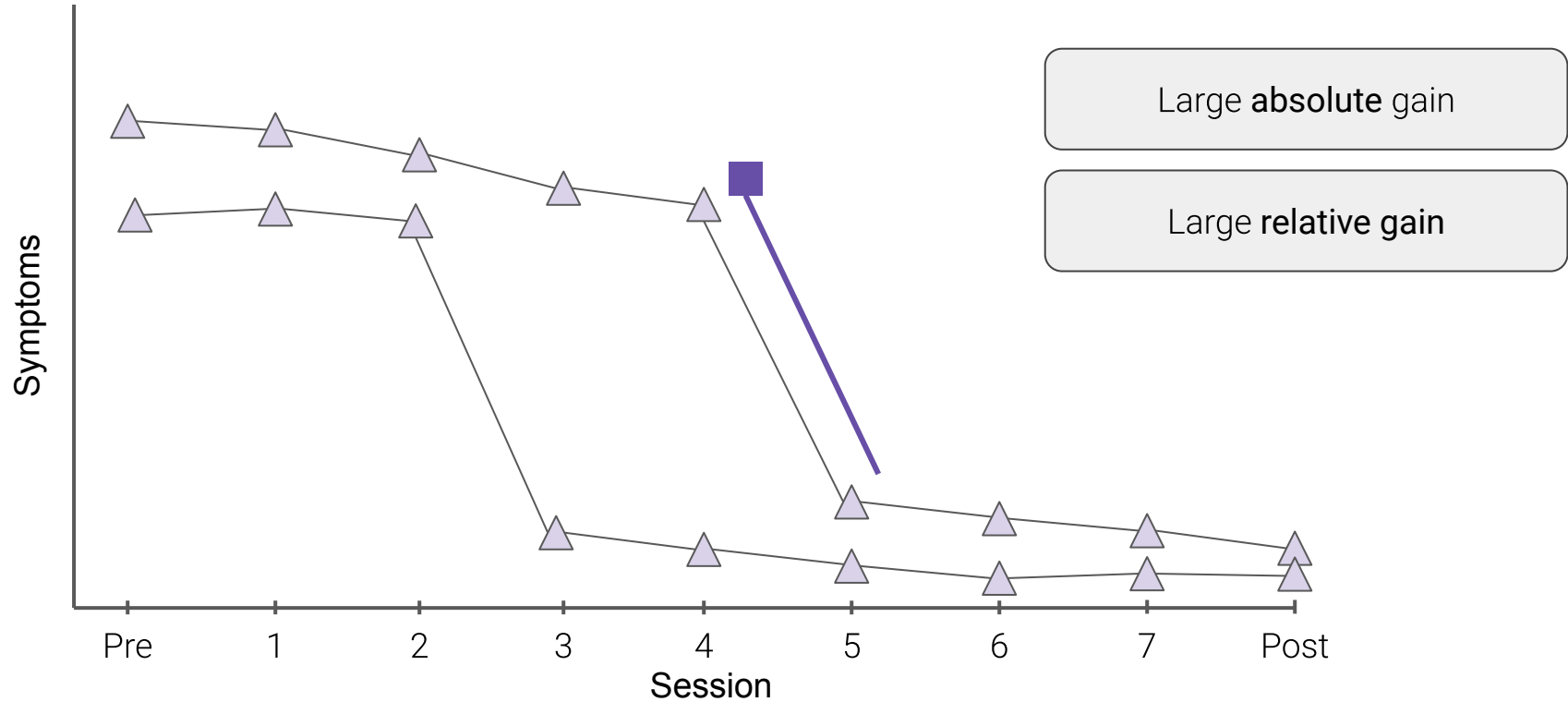


What are sudden gains?



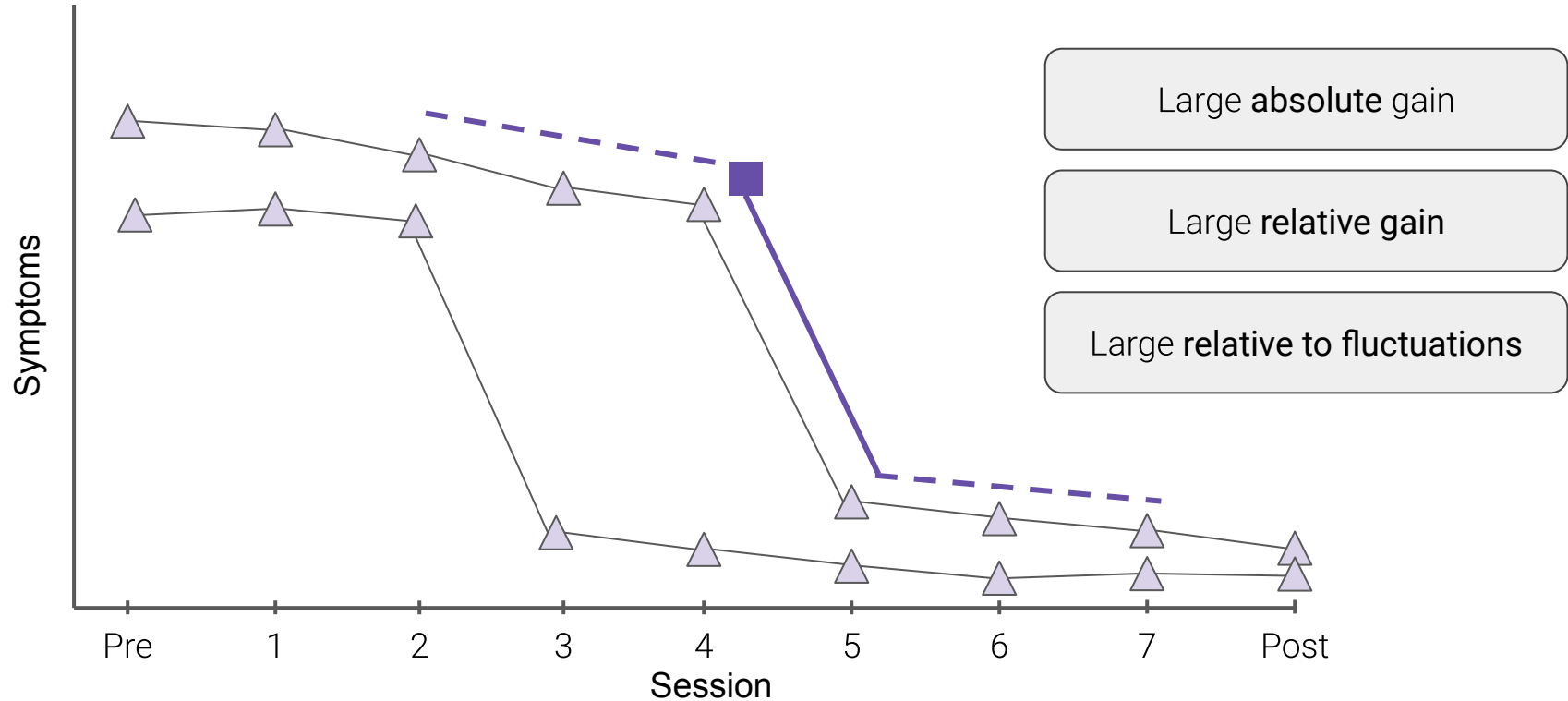
* Definition based on Tang & DeRubeis (1999).

What are sudden gains?



* Definition based on Tang & DeRubeis (1999).

What are sudden gains?



* Definition based on Tang & DeRubeis (1999).

For looooooooooooooops

Large **absolute** gain

Large **relative** gain

id	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10
1	35	35	37	20	24	21	24	21	11	11
2	29	26	27	21	20	6	4	5	6	3
3	NA	35	38	35	36	37	NA	22	11	8
4	33	NA	23	21	19	6	5	NA	NA	9

For loooooooooooooops

Large **absolute** gain

Large **relative** gain

Large **relative to fluctuations**

id	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10
1	35	35	37	20	24	21	24	21	11	11
2	29	26	27	21	20	6	4	5	6	3
3	NA	35	38	35	36	37	NA	22	11	8
4	33	NA	23	21	19	6	5	NA	NA	9

For looooooooooooooops

Large **absolute** gain

Large **relative** gain

Large **relative to fluctuations**

id	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10
1	35	35	37	20	24	21	24	21	11	11
2	29	26	27	21	20	6	4	5	6	3
3	NA	35	38	35	36	37	NA	22	11	8
4	33	NA	23	21	19	6	5	NA	NA	9

'Extract scores' around gains

id	n	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10
1	8	35	35	37	20	24	21	24	21	11	11
2	5	29	26	27	21	20	6	4	5	6	3
3	3	NA	35	38	35	36	37	NA	22	11	8
4	5	33	NA	23	21	19	6	5	NA	NA	9



What are the three values
before and after time point n?

id	2n	1n	n	n1	n2	n3
1	21	24	21	11	11	NA
2	27	21	20	6	4	5
3	37	NA	22	11	8	NA
4	23	21	19	6	5	NA

'Extract scores' around gains

id	n	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10
1	8	35	35	37	20	24	21	24	21	11	11
2	5	29	26	27	21	20	6	4	5	6	3
3	3	NA	35	38	35	36	37	NA	22	11	8
4	5	33	NA	23	21	19	6	5	NA	NA	9



What are the three values
before and after time point n?

id	2n	1n	n	n1	n2	n3
1	21	24	21	11	11	NA
2	27	21	20	6	4	5
3	37	NA	22	11	8	NA
4	23	21	19	6	5	NA

'Extract scores' around gains

id	n	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10
1	8	35	35	37	20	24	21	24	21	11	11
2	5	29	26	27	21	20	6	4	5	6	3
3	3	NA	35	38	35	36	37	NA	22	11	8
4	5	33	NA	23	21	19	6	5	NA	NA	9



What are the three values
before and after time point n?

id	2n	1n	n	n1	n2	n3
1	21	24	21	11	11	NA
2	27	21	20	6	4	5
3	37	NA	22	11	8	NA
4	23	21	19	6	5	NA

'Extract scores' around gains

id	n	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10
1	8	35	35	37	20	24	21	24	21	11	11
2	5	29	26	27	21	20	6	4	5	6	3
3	3	NA	35	38	35	36	37	NA	22	11	8
4	5	33	NA	23	21	19	6	5	NA	NA	9



What are the three values
before and after time point n?

id	2n	1n	n	n1	n2	n3
1	21	24	21	11	11	NA
2	27	21	20	6	4	5
3	37	NA	22	11	8	NA
4	23	21	19	6	5	NA

Reproducibility and productivity

The screenshot displays the RStudio interface with the following components:

- Source Editor:** Contains an R script named 'suddengains-gif.R'. The script includes comments and code for loading the 'suddengains' package, viewing the 'sgdata' dataset, identifying sudden gains, and creating a dataset with one gain per person.
- Environment Panel:** Shows the 'Global Environment' with a single object 'sgdata' containing 43 observations and 31 variables.
- Console:** Displays the command prompt at the bottom of the window.
- Terminal:** Shows the current working directory as '~/Projects/suddengains/'.

```
1 # 1. Load package ----
2 library(suddengains)
3
4 # 2. View example dataset ----
5 sgdata
6
7 # 3. Identify sudden gains ----
8 identify_sg(data = sgdata,
9             sg_crit1_cutoff = 7,
10            id_var_name = "id",
11            sg_var_list = c("bdi_s1", "bdi_s2", "bdi_s3",
12                           "bdi_s4", "bdi_s5", "bdi_s6",
13                           "bdi_s7", "bdi_s8", "bdi_s9",
14                           "bdi_s10", "bdi_s11", "bdi_s12"))
15
16 # 4. Create dataset with one gain per person ----
17 byperson <- create_byperson(data = sgdata,
18                             sg_crit1_cutoff = 7,
19                             id_var_name = "id",
20                             tx_start_var_name = "bdi_s1".
21
2:1 1. Load package :
```

Environment: suddengains

Global Environment

Data

Object	Size
sgdata	43 obs. of 31 variables

Files | Plots | Packages | Help | Viewer

~/Projects/suddengains/

>


```
library(suddengains)
```

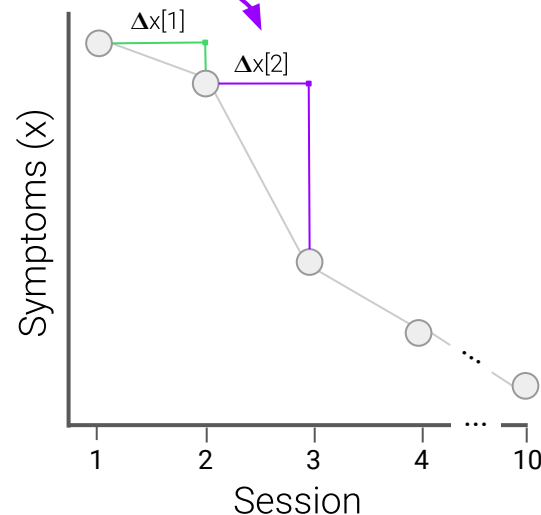
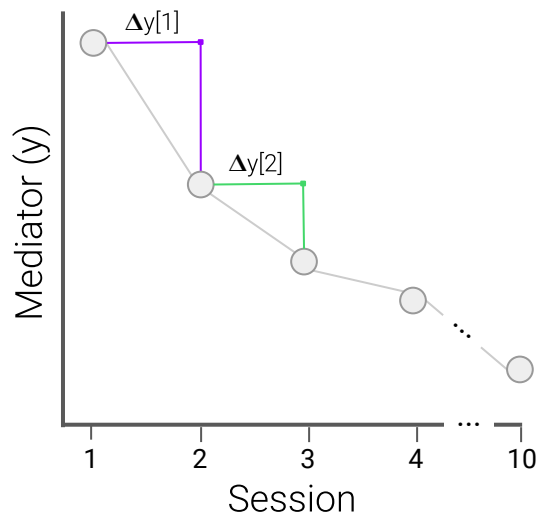
```
runApp(shinygains)
```


Modeling Change

Looking at change to change

Negative appraisals

PTSD symptoms

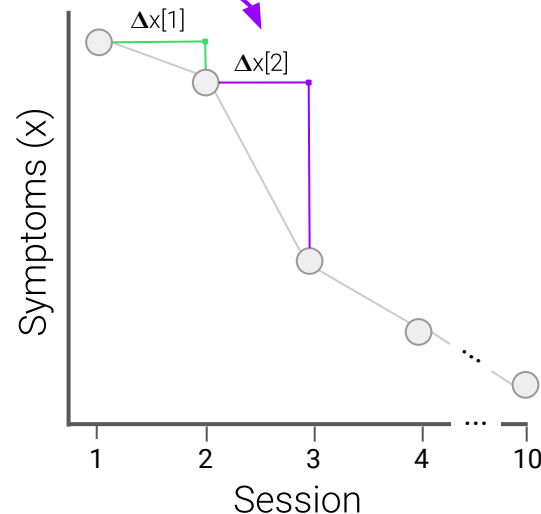
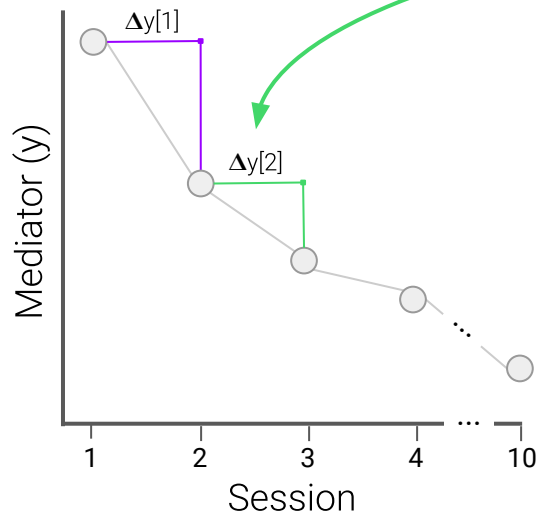



* Method implemented in the  package *lcs*m (Wiedemann, 2020) using *lavaan* (Rosseel, 2012). Overview of method in Grimm et al (2012).

Looking at change to change

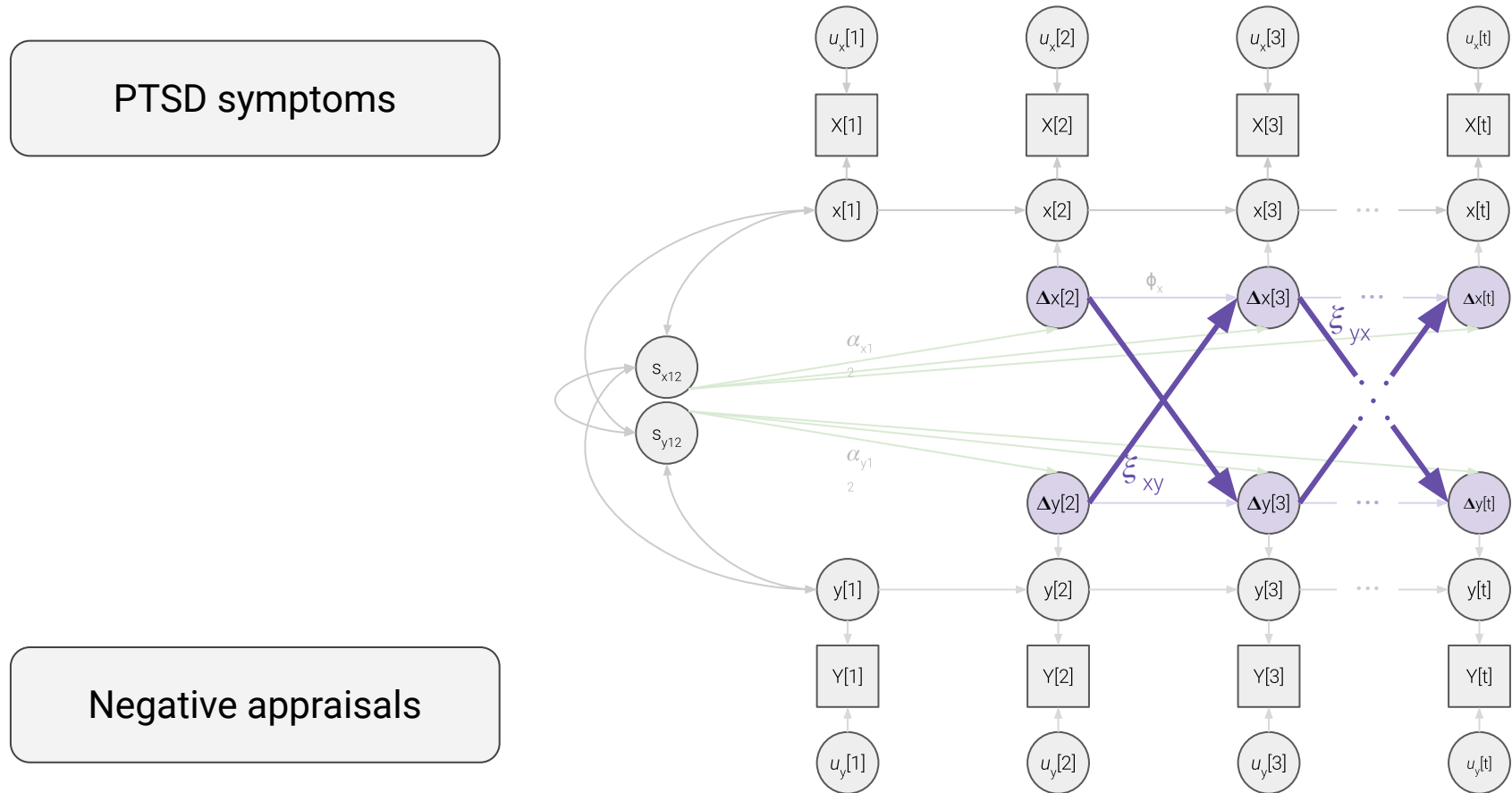
Negative appraisals

PTSD symptoms



* Method implemented in the  package *lscsm* (Wiedemann, 2020) using *lavaan* (Rosseel, 2012). Overview of method in Grimm et al (2012).

Bivariate latent change score modeling



Bivariate latent change score modeling

PTSD symptoms

Example lavaan syntax:

Regression:

$y \sim x_1 + x_2 + x_3$

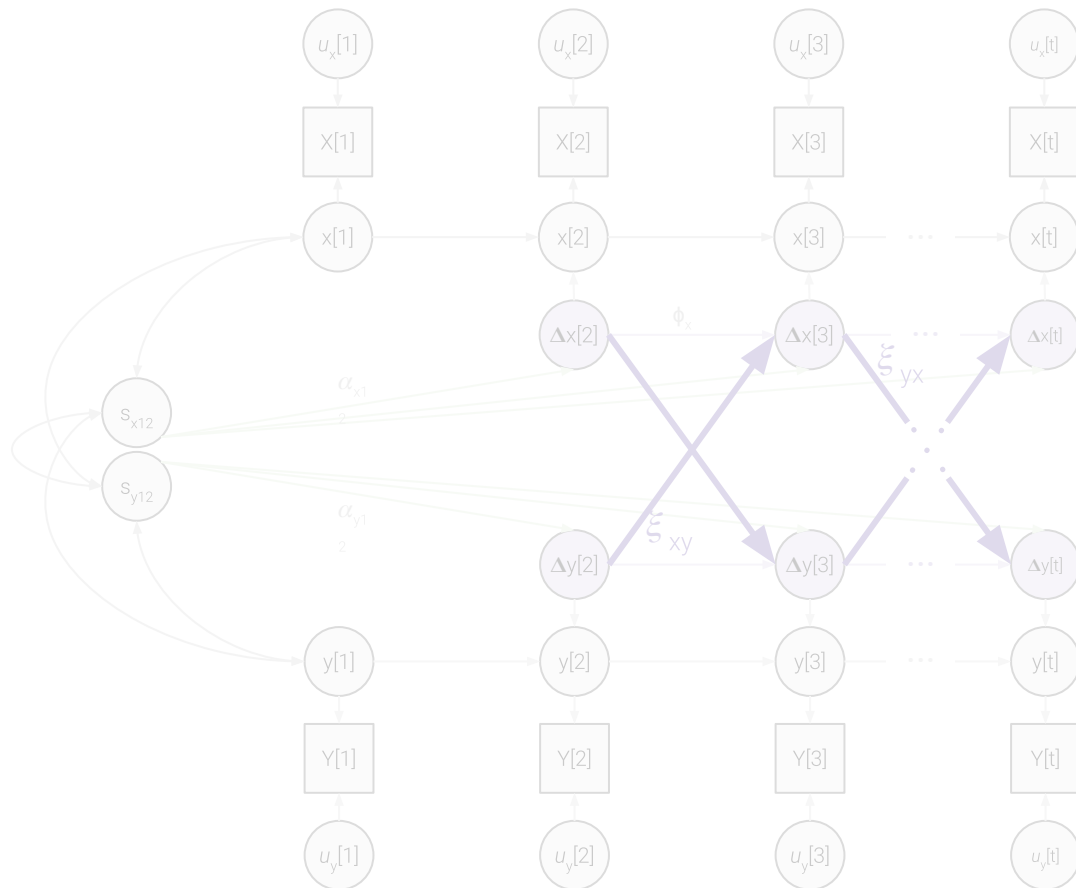
Latent Variable:

$11 \sim y_1 + y_2 + y_3$

Covariance:

$y_1 \sim y_2$

Negative appraisals



Tools to help understand latent change score modeling

shinychange

Overview

Generate lavaan Syntax ▾

Simulate Data ▾

Fit Model ▾

Options:

Data Characteristics

Select Parameters

Measurement Points:

3

Note: Number of repeated measurement points.

Variable Name:

x

Note: Variable name to be used for generating lavaan syntax, changes wont show on the path diagram. Variable name should start with a letter.

Results:

lavaan Syntax

Path Diagram

Note: lavaan syntax for the selected data characteristics and model parameters. This syntax includes comments describing the different sections of the model and can be modified by hand. Modified syntax could be used in the 'model' argument of functions from the lavaan package. Observed scores in the syntax are the variable name followed by a number indicating the measurement point. Latent true scores have the prefix 'l' (for latent) followed by the variable name of the observed score. Change scores have the prefix 'd' (for delta) followed by the variable name of the observed score.

```
# Specify latent true scores
lx1 =~ 1 * x1
lx2 =~ 1 * x2
lx3 =~ 1 * x3
# Specify mean of latent true scores
lx1 ~ gamma_lx1 * 1
lx2 ~ 0 * 1
lx3 ~ 0 * 1
# Specify variance of latent true scores
lx1 ~~ sigma2_lx1 * lx1
lx2 ~~ 0 * lx2
lx3 ~~ 0 * lx3
# Specify intercept of observed scores
x1 ~ 0 * 1
x2 ~ 0 * 1
x3 ~ 0 * 1
# Specify variance of observed scores
x1 ~~ sigma2_ux * x1
x2 ~~ sigma2_ux * x2
x3 ~~ sigma2_ux * x3
# Specify autoregressions of latent variables
lx2 ~ 1 * lx1
lx3 ~ 1 * lx2
# Specify latent change scores
```

Also see Kievit et al. (2018), Grimm et al (2017), and Ghisletta et al (2012).

```
library(lcsm)
```

```
runApp(shinychange)
```


Code written by some stranger on the internet is always perfect



Taking on Needless Dependencies

Fragile Development Guide

O RLY?

@ThePracticalDev

Software implementations are progressing at an increasingly fast pace*

Solutions that might fix the problem without breaking anything



Essential

Hoping This Works

O RLY?

@ThePracticalDev

* Reproducibility and Replicability in a Fast-Paced Methodological World by Epskamp ([2019](#)).