# Full Bayesian hierarchical model with latent dynamic structure for stress reactivity and five EMA PID-5 trait domains

Corrado Caudek

## Dati PID-5 e ESI-BF (baseline)

Le misure "basali" corrispondenti ai 5 domini del PID-5 sono state calcolate **escludendo** i 15 item che vengono usati nelle notifiche EMA.

```
# Importa i punteggi ESI-BF (fattore di rischio psicologico globale)
# e tiene solo un record per soggetto
esi_bf <- rio::import(
  here::here(
    "data",
    "processed",
    "esi_bf.csv"
  )
) |>
  dplyr::distinct(user_id, .keep_all = TRUE) |> # Keep only distinct user_id
  dplyr::select(user_id, esi_bf) # Select relevant columns

# Importa i punteggi del PID-5, escludendo gli item usati nell'EMA
pid5 <- rio::import(
  here::here(
    "data",
    "processed",
    "pid5.csv"
  )
) |>
  dplyr::distinct(user_id, .keep_all = TRUE) |>  # Keep only distinct user_id
  dplyr::select(user_id, starts_with("domain_")) # Select domain variables

# Merge dei dati basali in un unico dataframe
df <- left_join(esi_bf, pid5, by = "user_id")
```

## Pulizia iniziale dei dati

```
# Elimina soggetti con risposte non affidabili (careless responding)
user_id_with_careless_responding <- c(
  "ma_se_2005_11_14_490",
  "reve20041021036",
  "di_ma_2005_10_20_756",
  "pa_sc_2005_09_10_468",
  "il_re_2006_01_18_645",
  "so_ma_2003_10_13_804",
```

```r
    "lo_ca_2005_05_07_05_437",
    "va_ma_2005_05_31_567",
    "no_un_2005_06_29_880",
    "an_bo_1988_08_24_166",
    "st_ma_2004_04_21_426",
    "an_st_2005_10_16_052",
    "vi_de_2002_12_30_067",
    "gi_ru_2005_03_08_033",
    "al_mi_2005_03_05_844",
    "la_ma_2006_01_31_787",
    "gi_lo_2004_06_27_237",
    "ch_bi_2001_01_28_407",
    "al_pe_2001_04_20_079",
    "le_de_2003_09_05_067",
    "fe_gr_2002_02_19_434",
    "ma_ba_2002_09_09_052",
    "ca_gi_2003_09_16_737",
    "an_to_2003_08_06_114",
    "al_se_2003_07_28_277",
    "ja_tr_2002_10_06_487",
    "el_ci_2002_02_15_057",
    "se_ti_2000_03_04_975",
    "co_ga_2003_10_29_614",
    "al_ba_2003_18_07_905",
    "bi_ro_2003_09_07_934",
    "an_va_2004_04_08_527",
    "ev_cr_2003_01_27_573"
)

# Filter out users with careless responses
df1 <- df[!(df$user_id %in% user_id_with_careless_responding), ]
```

### Caricamento dati EMA

```r
# Legge i dati grezzi EMA e li unisce ai dati baseline
ema_raw <- readRDS(
  here::here(
    "data",
    "raw",
    "ema",
    "ema_data_scoring.RDS"
  )
) |>
  dplyr::rename(
    user_id = subj_code
  )

# Merge EMA data with filtered main data
df2 <- left_join(df1, ema_raw, by = "user_id")

# Verify number of unique users
length(unique(df2$user_id))
```

2

```
[1] 429

# Conta quante risposte EMA ha fornito ciascun soggetto
user_counts <- df2 %>%
  group_by(user_id) %>%
  summarise(n_responses = n()) %>%
  ungroup()

valid_users <- user_counts %>%
  filter(n_responses >= 10) %>%
  pull(user_id)

# Tiene solo i soggetti con almeno 10 risposte EMA
df2 <- df2 %>%
  dplyr::filter(user_id %in% valid_users)

length(unique(df2$user_id))

[1] 379

glimpse(df2)

Rows: 10,024
Columns: 89
$ user_id                <chr> "an_ra_2004_06_07_770", "an_ra_2004_06_07_77~
$ esi_bf                 <int> 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8,~
$ domain_negative_affect <int> 35, 35, 35, 35, 35, 35, 35, 35, 35, 35, 35, ~
$ domain_detachment      <int> 18, 18, 18, 18, 18, 18, 18, 18, 18, 18, 18, ~
$ domain_antagonism      <int> 18, 18, 18, 18, 18, 18, 18, 18, 18, 18, 18, ~
$ domain_disinhibition   <int> 23, 23, 23, 23, 23, 23, 23, 23, 23, 23, 23, ~
$ domain_psychoticism    <int> 26, 26, 26, 26, 26, 26, 26, 26, 26, 26, 26, ~
$ happy                  <dbl> 61, 100, 77, 100, 35, 15, 29, 46, 74, 81, 22~
$ sad                    <dbl> 16, 0, 10, 10, 65, 62, 39, 52, 24, 10, 66, 7~
$ satisfied              <dbl> 56, 100, 84, 91, 42, 20, 36, 19, 76, 88, 13,~
$ angry                  <dbl> 0, 0, 0, 0, 19, 26, 50, 67, 7, 14, 22, 50, 8~
$ pid5_13                <dbl> 0, 1, 2, 1, 0, 0, 0, 0, 0, 0, 0, 2, 0, 0, 0,~
$ pid5_15                <dbl> 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0,~
$ pid5_11                <dbl> 2, 2, 2, 0, 2, 2, 0, 0, 0, 0, 2, 2, 1, 0, 0,~
$ pid5_3                 <dbl> 1, 1, 3, 2, 2, 0, 0, 0, 0, 0, 2, 2, 2, 2, 0,~
$ pid5_2                 <dbl> 2, 2, 1, 0, 2, 2, 0, 3, 0, 0, 0, 1, 2, 0, 0,~
$ pid5_7                 <dbl> 2, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,~
$ pid5_14                <dbl> 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,~
$ pid5_6                 <dbl> 2, 3, 2, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 2,~
$ pid5_4                 <dbl> 2, 3, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0,~
$ pid5_12                <dbl> 1, 2, 1, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0,~
$ pid5_1                 <dbl> 2, 2, 0, 0, 1, 1, 0, 2, 0, 0, 2, 2, 1, 0, 0,~
$ pid5_9                 <dbl> 1, 3, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,~
$ pid5_5                 <dbl> 2, 2, 3, 0, 0, 0, 0, 0, 0, 0, 2, 1, 2, 2, 0,~
$ pid5_8                 <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,~
$ pid5_10                <dbl> 1, 3, 1, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0,~
$ tripm_1                <dbl> 2, 1, 4, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,~
$ tripm_3                <dbl> 3, 1, 1, 1, 1, 1, 4, 1, 1, 1, 1, 1, 1, 1, 1,~
$ tripm_2                <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,~
$ tripm_4                <dbl> 4, 2, 3, 3, 3, 3, 3, 3, 3, 4, 3, 3, 4, 3, 3,~
```

```
$ dass21_2                   <dbl> 0, NA, 2, NA, 0, NA, NA, 0, 0, NA, 2, NA, 0,~
$ dass21_5                   <dbl> 1, NA, 0, NA, 0, NA, NA, 0, 0, NA, 0, NA, 0,~
$ dass21_3                   <dbl> 1, NA, 0, NA, 0, NA, NA, 0, 0, NA, 0, NA, 0,~
$ dass21_4                   <dbl> 1, NA, 0, NA, 0, NA, NA, 2, 0, NA, 0, NA, 0,~
$ dass21_6                   <dbl> 0, NA, 0, NA, 0, NA, NA, 0, 0, NA, 0, NA, 0,~
$ dass21_1                   <dbl> 0, NA, 2, NA, 0, NA, NA, 1, 0, NA, 0, NA, 0,~
$ scs3_pos                   <dbl> 3, NA, 3, NA, 3, NA, NA, 3, -3, NA, -3, NA, ~
$ scs5_neg                   <dbl> -3, NA, -1, NA, -2, NA, NA, -2, -3, NA, -2, ~
$ scs7_pos                   <dbl> 1, NA, 2, NA, 2, NA, NA, 2, 1, NA, 2, NA, -1~
$ scs4_neg                   <dbl> 2, NA, 2, NA, -3, NA, NA, -3, -2, NA, 2, NA,~
$ scs2_neg                   <dbl> -1, NA, -2, NA, -3, NA, NA, -3, -3, NA, 2, N~
$ scs1_pos                   <dbl> -1, NA, 3, NA, 2, NA, NA, 2, -3, NA, 2, NA, ~
$ scs8_neg                   <dbl> 1, NA, 2, NA, -2, NA, NA, -3, -1, NA, 1, NA,~
$ scs6_pos                   <dbl> 1, NA, 2, NA, 2, NA, NA, 2, 2, NA, 2, NA, -3~
$ vq_2                       <dbl> NA, 0, NA, 0, NA, 0, 0, 0, 0, 0, NA, 0, NA, ~
$ vq_1                       <dbl> NA, 0, NA, 0, NA, 0, 0, 0, 0, 0, NA, 0, NA, ~
$ vq_4                       <dbl> NA, 0, NA, 0, NA, 0, 0, 0, 0, 0, NA, 0, NA, ~
$ vq_3                       <dbl> NA, 0, NA, 0, NA, 0, 0, 0, 0, 0, NA, 0, NA, ~
$ cope_nvi_2                 <dbl> NA, 4, NA, 2, NA, 2, 1, 3, 1, 2, NA, 1, NA, ~
$ cope_nvi_7                 <dbl> NA, 4, NA, 2, NA, 3, 4, 4, 3, 4, NA, 2, NA, ~
$ cope_nvi_5                 <dbl> NA, 4, NA, 3, NA, 4, 3, 4, 3, 4, NA, 4, NA, ~
$ cope_nvi_8                 <dbl> NA, 4, NA, 3, NA, 4, 3, 4, 4, 4, NA, 3, NA, ~
$ cope_nvi_9                 <dbl> NA, 1, NA, 1, NA, 1, 1, 1, 1, 1, NA, 1, NA, ~
$ cope_nvi_10                <dbl> NA, 1, NA, 1, NA, 1, 2, 1, 1, 1, NA, 4, NA, ~
$ cope_nvi_1                 <dbl> NA, 2, NA, 1, NA, 1, 1, 1, 1, 1, NA, 1, NA, ~
$ cope_nvi_3                 <dbl> NA, 3, NA, 3, NA, 3, 3, 4, 4, 4, NA, 2, NA, ~
$ cope_nvi_4                 <dbl> NA, 4, NA, 3, NA, 3, 3, 4, 4, 4, NA, 2, NA, ~
$ cope_nvi_6                 <dbl> NA, 4, NA, 3, NA, 4, 3, 4, 4, 4, NA, 4, NA, ~
$ day                        <date> 2025-03-22, 2025-03-26, 2025-03-29, 2025-04~
$ hour                       <int> 9, 18, 9, 18, 9, 18, 9, 21, 17, 18, 18, 11, ~
$ calendar_day               <dbl> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 1~
$ bysubj_day                 <int> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 1~
$ context_quality            <dbl> 1, 2, 0, 2, 0, 0, -1, -1, 2, 2, -1, 0, -1, 1~
$ context_control            <dbl> 2, 4, 3, 5, 3, 2, 2, 2, 4, 5, 2, 2, 1, 4, 5,~
$ context_support            <dbl> 2, 2, 4, 3, 3, 2, 2, 5, 3, 5, 3, 3, 1, 3, 5,~
$ context_threat             <dbl> 1, 0, 1, 0, 1, 2, 1, 2, 0, 0, 2, 2, 2, 0, 1,~
$ pid5_sum                   <dbl> 20, 25, 17, 8, 10, 9, 3, 8, 3, 3, 12, 15, 11~
$ pid5_negative_affectivity  <dbl> 5, 5, 4, 2, 5, 3, 0, 5, 0, 0, 4, 5, 5, 2, 0,~
$ pid5_detachment            <dbl> 6, 8, 5, 3, 3, 4, 3, 3, 3, 3, 5, 4, 5, 5, 2,~
$ pid5_antagonism            <dbl> 3, 3, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,~
$ pid5_disinhibition         <dbl> 4, 7, 4, 2, 2, 2, 0, 0, 0, 0, 3, 3, 1, 0, 0,~
$ pid5_psychoticism          <dbl> 2, 2, 3, 1, 0, 0, 0, 0, 0, 0, 0, 3, 0, 0, 0,~
$ ipv_sum                    <dbl> NA, 0, NA, 0, NA, 0, 0, 0, 0, 0, NA, 0, NA, ~
$ tripm_4_rev                <dbl> 1, 3, 2, 2, 2, 2, 2, 2, 2, 1, 2, 2, 1, 2, 2,~
$ tripm_sum                  <dbl> 10, 5, 9, 6, 7, 6, 9, 6, 6, 7, 6, 6, 7, 6, 6~
$ tripm_boldness             <dbl> 3, 2, 5, 2, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,~
$ tripm_meanness             <dbl> 4, 4, 3, 3, 3, 3, 6, 3, 3, 2, 3, 3, 2, 3, 3,~
$ dass_sum                   <dbl> 3, NA, 4, NA, 0, NA, NA, 3, 0, NA, 2, NA, 0,~
$ dass_stress                <dbl> 0, NA, 4, NA, 0, NA, NA, 1, 0, NA, 2, NA, 0,~
$ dass_depression            <dbl> 2, NA, 0, NA, 0, NA, NA, 2, 0, NA, 0, NA, 0,~
$ dass_anxiety               <dbl> 1, NA, 0, NA, 0, NA, NA, 0, 0, NA, 0, NA, 0,~
```

4

```
$ cope_10_rev              <dbl> NA, 4, NA, 4, NA, 4, 3, 4, 4, 4, NA, 1, NA, ~
$ cope_avoid              <dbl> NA, 6, NA, 3, NA, 3, 2, 4, 2, 3, NA, 2, NA, ~
$ cope_prob_or            <dbl> NA, 7, NA, 6, NA, 6, 6, 8, 8, 8, NA, 4, NA, ~
$ cope_social_support     <dbl> NA, 8, NA, 6, NA, 8, 6, 8, 7, 8, NA, 8, NA, ~
$ cope_positive_att       <dbl> NA, 8, NA, 5, NA, 7, 7, 8, 7, 8, NA, 5, NA, ~
$ cope_trascendent_or     <dbl> NA, 5, NA, 5, NA, 5, 4, 5, 5, 5, NA, 2, NA, ~
$ cs_pos                  <dbl> 4, NA, 10, NA, 9, NA, NA, 9, -3, NA, 3, NA, ~
$ ucs_neg                 <dbl> -1, NA, 1, NA, -10, NA, NA, -11, -9, NA, 3, ~
```

## Costruzione della variabile 'exam_period'

```r
# Importa i dati delle date d'esame per ogni soggetto
exam_data <- rio::import(
  here::here(
    "data", "raw", "exam_data.xlsx"
  )
) %>%
  dplyr::select(user_id = subj_code, exam_day = date, course, sex)

# Pulisce le date e ne estrae la componente valida
exam_data_clean <- exam_data %>%
  mutate(
    # Estrae solo la parte che somiglia a una data
    exam_day_extracted = str_extract(exam_day, "\\b\\d{1,2}[-_\\./]\\d{1,2}[-_\\./]\\d{2,4}`
    # Sostituisce separatori misti con "/"
    exam_day_clean = str_replace_all(exam_day_extracted, "[-_\\.]", "/"),
    # Converte in formato Date
    exam_day_date = suppressWarnings(dmy(exam_day_clean))
  )

# Sostituisce le date per il corso Psicometria con quelle ufficiali
psicom_dates <- tibble(
  exam_day_date = as.Date(c("2025-04-15", "2025-05-22", "2025-05-26"))
)

exam_data_psicom <- exam_data_clean %>%
  dplyr::filter(course == "Psicometria") %>%
  dplyr::distinct(user_id, course, sex) %>%
  tidyr::crossing(psicom_dates)

# Mantiene dati validi per altri corsi (solo se la data è valida)
exam_data_other <- exam_data_clean %>%
  dplyr::filter(course != "Psicometria", !is.na(exam_day_date)) %>%
  dplyr::select(user_id, course, sex, exam_day_date)

# Unisce i due
exam_data_fixed <- bind_rows(exam_data_psicom, exam_data_other)

# Estrae le date, corregge il formato, e crea variabili per: giorno prima,
# giorno dopo, giorno esame
exam_data_long <- exam_data_fixed %>%
  mutate(
    pre_exam_1 = exam_day_date - days(1),
    pre_exam_2 = exam_day_date - days(2),
    post_exam = exam_day_date + days(1)
```

```r
  ) %>%
  dplyr::select(user_id, pre_exam_1, pre_exam_2, exam_day = exam_day_date, post_exam) %>%
  pivot_longer(
    cols = c(pre_exam_1, pre_exam_2, exam_day, post_exam),
    names_to = "exam_period_raw",
    values_to = "exam_date"
  ) %>%
  mutate(
    exam_period = case_when(
      exam_period_raw %in% c("pre_exam_1", "pre_exam_2") ~ "pre_exam",
      exam_period_raw %in% c("exam_day", "post_exam") ~ "post_exam"
    )
  )

# Unisce le informazioni EMA e d'esame in un unico dataframe
df_exam_tagged <- df2 %>%
  left_join(exam_data_long,
            by = c("user_id" = "user_id", "day" = "exam_date")) %>%
  mutate(
    exam_period = coalesce(exam_period, "baseline")  # se NA, è "baseline"
  )

table(df_exam_tagged$exam_period)
```

```
 baseline post_exam  pre_exam
     8879       602       543
```

### Variabile binaria `under_stress`

```r
df_exam_tagged <- df_exam_tagged %>%
  mutate(under_stress = if_else(exam_period %in% c("pre_exam", "post_exam"), 1, 0))
```

In questo script la variabile `under_stress` non viene usata.

### Calcolo affetto negativo istantaneo

```r
# Costruisce una misura media dell'affetto negativo momentaneo

# Seleziona solo le colonne rilevanti (per velocità)
items <- c("sad", "angry", "happy", "satisfied")

# Imputazione dei missing su variabili di umore
imputed <- mice(df_exam_tagged[, items], m = 1, maxit = 10, seed = 123)
```

```
 iter imp variable
  1   1  sad  angry  happy  satisfied
  2   1  sad  angry  happy  satisfied
  3   1  sad  angry  happy  satisfied
  4   1  sad  angry  happy  satisfied
  5   1  sad  angry  happy  satisfied
  6   1  sad  angry  happy  satisfied
  7   1  sad  angry  happy  satisfied
```

```
8   1  sad  angry  happy  satisfied
9   1  sad  angry  happy  satisfied
10   1  sad  angry  happy  satisfied
```

```r
# Estrai il dataset imputato e sostituisci le colonne originali
df2_imputed <- complete(imputed)
df2[, items] <- df2_imputed[, items]

# Calcolo della media invertendo happy/satisfied
df_exam_tagged <- df_exam_tagged %>%
  mutate(
    happy_reversed = 100 - happy, # Scala 0-100
    satisfied_reversed = 100 - satisfied,
    neg_aff_ema = rowMeans(
      cbind(sad, angry, happy_reversed, satisfied_reversed),
      na.rm = TRUE
    )
  )
```

### Variabili `delta_ucs` e `delta_mood`

Se i valori di pre-esame sono più alti della baseline e i valori di post-esame sono più bassi della baseline, allora la formula (post - pre) / baseline potrebbe effettivamente catturare meglio l'effetto complessivo dello stress.

- (post - pre) cattura la direzione del cambiamento (se i valori stanno diminuendo o aumentando);
- La divisione per baseline normalizza il cambiamento rispetto al livello basale, rendendo l'effetto comparabile tra soggetti con diversi livelli di partenza.

```r
# df_stress_effects <- df_exam_tagged %>%
#   dplyr::filter(exam_period %in% c("baseline", "post_exam")) %>%
#   group_by(user_id, exam_period) %>%
#   summarise(
#     ucs = mean(ucs_neg, na.rm = TRUE),
#     mood = mean(neg_aff_ema, na.rm = TRUE),
#     .groups = "drop"
#   ) %>%
#   pivot_wider(names_from = exam_period, values_from = c(ucs, mood)) %>%
#   mutate(
#     delta_ucs = ucs_post_exam - ucs_baseline,
#     delta_mood = mood_post_exam - mood_baseline
#   )

df_stress_effects <- df_exam_tagged %>%
  dplyr::filter(exam_period %in% c("baseline", "pre_exam", "post_exam")) %>%
  group_by(user_id, exam_period) %>%
  summarise(
    ucs = mean(ucs_neg, na.rm = TRUE),
    mood = mean(neg_aff_ema, na.rm = TRUE),
    .groups = "drop"
  ) %>%
  pivot_wider(names_from = exam_period, values_from = c(ucs, mood)) %>%
```

```
  mutate(
    delta_ucs = (ucs_post_exam - ucs_pre_exam) / ucs_baseline,
    delta_mood = (mood_post_exam - mood_pre_exam) / mood_baseline
  )
```

Qui soltanto descrivo il pattern di cambiamento (pre/post) rispetto alla baseline.

```
mod1 <- brm(
  ucs_neg ~ exam_period + (1 + exam_period | user_id),
  data = df_exam_tagged,
  family = gaussian(),
  backend = "cmdstanr"
)
```

```
pp_check(mod1)
```

Using 10 posterior draws for ppc type 'dens_overlay' by default.
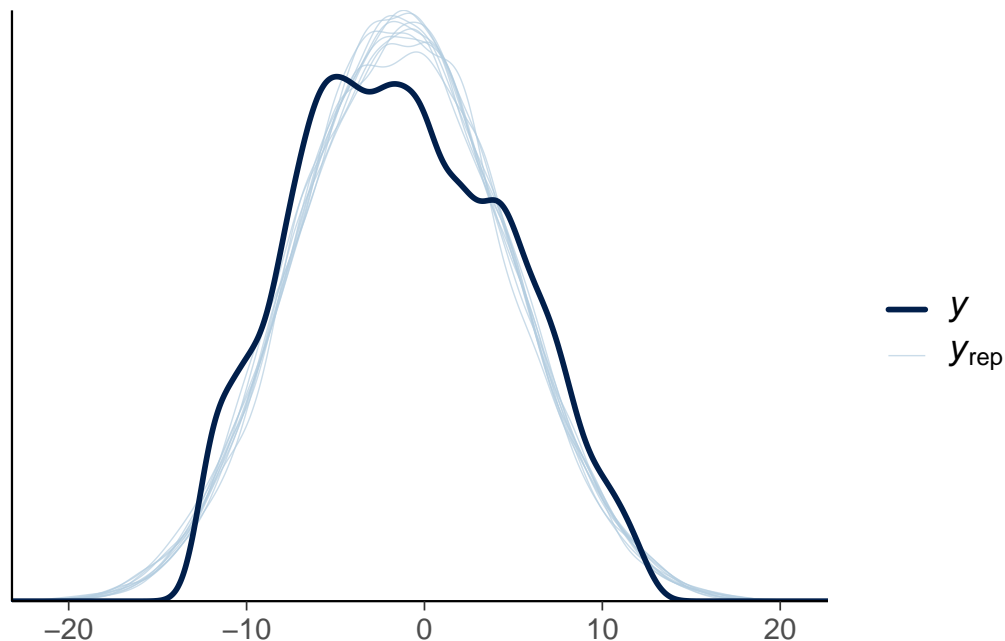


```
print(mod1)
```

```
 Family: gaussian
  Links: mu = identity; sigma = identity
Formula: ucs_neg ~ exam_period + (1 + exam_period | user_id)
   Data: df_exam_tagged (Number of observations: 6229)
  Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup draws = 4000
```

Multilevel Hyperparameters:
~user_id (Number of levels: 379)

| | Estimate | Est.Error | l-95% CI |
|---|---|---|---|
| sd(Intercept) | 4.32 | 0.17 | 3.99 |
| sd(exam_periodpost_exam) | 0.77 | 0.45 | 0.04 |
| sd(exam_periodpre_exam) | 0.76 | 0.45 | 0.04 |
| cor(Intercept,exam_periodpost_exam) | -0.16 | 0.30 | -0.74 |
| cor(Intercept,exam_periodpre_exam) | 0.17 | 0.31 | -0.49 |

```
cor(exam_periodpost_exam,exam_periodpre_exam)        0.29        0.48      -0.77
                                              u-95% CI Rhat Bulk_ESS Tail_ESS
sd(Intercept)                                     4.67 1.01      356     1090
sd(exam_periodpost_exam)                          1.65 1.01      403     1241
sd(exam_periodpre_exam)                           1.67 1.02      460     1389
cor(Intercept,exam_periodpost_exam)               0.53 1.00     3631     1586
cor(Intercept,exam_periodpre_exam)                0.80 1.01     3575     2061
cor(exam_periodpost_exam,exam_periodpre_exam)     0.94 1.00      660     1940

Regression Coefficients:
                     Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS
Intercept               -1.23      0.23    -1.67    -0.79 1.02      255
exam_periodpost_exam    -0.79      0.20    -1.18    -0.39 1.00     5065
exam_periodpre_exam      0.76      0.23     0.33     1.21 1.00     5578
                     Tail_ESS
Intercept                 440
exam_periodpost_exam     3080
exam_periodpre_exam      3249

Further Distributional Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma     3.88      0.04     3.81     3.95 1.00     4607     2880

Draws were sampled using sample(hmc). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```
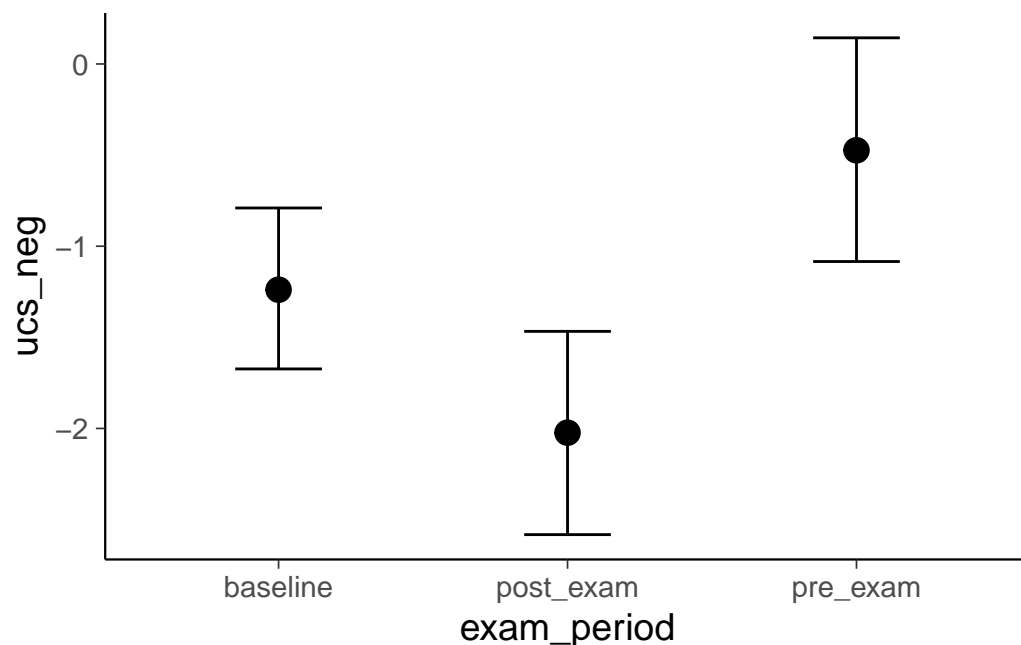
conditional_effects(mod1, "exam_period")



```
tapply(df_exam_tagged$ucs_neg, df_exam_tagged$exam_period, mean, na.rm = TRUE)

  baseline  post_exam   pre_exam
-1.2906692 -1.7577519  0.3602941
```

```r
tapply(df_exam_tagged$cs_pos, df_exam_tagged$exam_period, mean, na.rm = TRUE)
```

```
 baseline post_exam  pre_exam
1.9788878 1.5717054 0.7696078
```

```r
tapply(df_exam_tagged$neg_aff_ema, df_exam_tagged$exam_period, mean, na.rm = TRUE)
```

```
 baseline post_exam  pre_exam
 37.19474  33.25955  42.62661
```

```r
df_with_delta <- left_join(df_exam_tagged, df_stress_effects,  by = "user_id")

mod2 <- brm(
  delta_ucs ~ domain_negative_affect + domain_detachment + domain_antagonism +
    domain_disinhibition + domain_psychoticism +
    pid5_negative_affectivity + pid5_detachment + pid5_antagonism +
    pid5_disinhibition + pid5_psychoticism + (1 | user_id/day),
  data = df_with_delta,
  family = gaussian(),
  backend = "cmdstanr",
  algorithm = "meanfield",
  seed = 123
)

summary(mod2)
```

```
 Family: gaussian
  Links: mu = identity; sigma = identity
Formula: delta_ucs ~ domain_negative_affect + domain_detachment + domain_antagonism + domain
   Data: df_with_delta (Number of observations: 3397)
  Draws: 1 chains, each with iter = 1000; warmup = 0; thin = 1;
         total post-warmup draws = 1000


Multilevel Hyperparameters:
~user_id (Number of levels: 122)
              Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sd(Intercept)     2.72      0.01     2.71     2.73 1.01      914      699


~user_id:day (Number of levels: 3138)
              Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sd(Intercept)     0.01      0.01     0.00     0.02 1.00     1088      916


Regression Coefficients:
                          Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS
Intercept                     2.74      0.06     2.61     2.87 1.00      940
domain_negative_affect       -0.05      0.00    -0.05    -0.04 1.00      909
domain_detachment             0.03      0.00     0.02     0.03 1.00      880
domain_antagonism            -0.05      0.00    -0.06    -0.05 1.00     1027
domain_disinhibition         -0.03      0.00    -0.03    -0.03 1.00      947
domain_psychoticism          -0.01      0.00    -0.01    -0.00 1.00     1003
pid5_negative_affectivity    -0.00      0.01    -0.01     0.01 1.00      971
pid5_detachment              -0.00      0.00    -0.01     0.01 1.00      914
pid5_antagonism              -0.01      0.01    -0.02     0.00 1.00     1008
pid5_disinhibition           -0.00      0.01    -0.02     0.02 1.00      870
pid5_psychoticism             0.00      0.01    -0.01     0.02 1.00     1010
```

```
                         Tail_ESS
Intercept                      983
domain_negative_affect         912
domain_detachment              696
domain_antagonism             1058
domain_disinhibition           975
domain_psychoticism            906
pid5_negative_affectivity      875
pid5_detachment                845
pid5_antagonism                884
pid5_disinhibition             992
pid5_psychoticism              981


Further Distributional Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma     0.54      0.01     0.52     0.56 1.00      946      944


Draws were sampled using variational(meanfield).
```

## Reattività individuale allo stress (delta)

```r
# Calcola cambiamenti tra baseline e post_exam per UCS, CS, umore
df_stress_profile <- df_exam_tagged %>%
  dplyr::filter(exam_period %in% c("baseline", "pre_exam", "post_exam")) %>%
  group_by(user_id, exam_period) %>%
  summarise(
    ucs = mean(ucs_neg, na.rm = TRUE),
    cs  = mean(cs_pos, na.rm = TRUE),
    mood = mean(neg_aff_ema, na.rm = TRUE),
    .groups = "drop"
  ) %>%
  pivot_wider(names_from = exam_period, values_from = c(ucs, cs, mood)) %>%
  mutate(
    delta_ucs  = ucs_post_exam - ucs_baseline,
    delta_cs   = cs_post_exam  - cs_baseline,
    delta_mood = mood_post_exam - mood_baseline,

    # Opzionale: anche pre_exam
    delta_ucs_pre  = ucs_pre_exam - ucs_baseline,
    delta_cs_pre   = cs_pre_exam  - cs_baseline,
    delta_mood_pre = mood_pre_exam - mood_baseline
  )
```

Costruzione indice composito `stress_reactivity`.

```r
df_stress_profile <- df_stress_profile %>%
  mutate(
    z_delta_ucs  = scale(delta_ucs)[,1],
    z_delta_cs   = scale(-delta_cs)[,1],    # notare il "-" → calo della cs = peggioramento
    z_delta_mood = scale(delta_mood)[,1]
  ) %>%
  mutate(
    stress_reactivity = z_delta_ucs + z_delta_cs + z_delta_mood
```

```
)
hist(df_stress_profile$stress_reactivity)
```

## Histogram of df_stress_profile$stress_reactivity
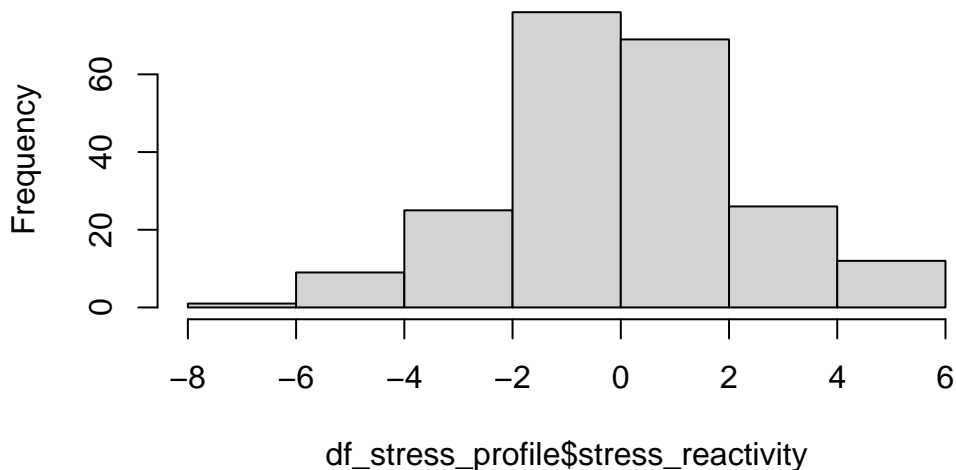


```
df_stress_profile <- df_stress_profile %>%
  mutate(
    ucs_baseline_z = scale(ucs_baseline)[,1],
    cs_baseline_z  = scale(cs_baseline)[,1],
    mood_baseline_z = scale(mood_baseline)[,1]
  )

fm <- lm(
  stress_reactivity ~ ucs_baseline_z + cs_baseline_z + mood_baseline_z,
  data = df_stress_profile)

# 1. Seleziona solo le colonne rilevanti e rimuovi i casi incompleti
df_for_pca <- df_stress_profile %>%
  dplyr::select(user_id, z_delta_ucs, z_delta_cs, z_delta_mood) %>%
  drop_na()

# 2. Applica PCA senza ulteriori scaling (le variabili sono già standardizzate)
react_pca <- prcomp(df_for_pca %>% dplyr::select(-user_id), scale. = FALSE)

# 3. Aggiungi il primo componente al dataframe
df_for_pca <- df_for_pca %>%
  mutate(stress_reactivity_pca = react_pca$x[, 1])

df_stress_profile <- df_stress_profile %>%
  left_join(df_for_pca %>% dplyr::select(user_id, stress_reactivity_pca), by = "user_id")

pid5_baseline <- df_exam_tagged |>
  dplyr::select(
    user_id, domain_negative_affect, domain_detachment, domain_antagonism,
    domain_disinhibition, domain_psychoticism
  ) |>
```

```r
  dplyr::distinct(user_id, .keep_all = TRUE)

df <- left_join(df_stress_profile, pid5_baseline, by = "user_id")
```

## Regressioni lineari: baseline e tratti EMA

```r
# Regressione su tratti basali PID-5
model_pid5 <- lm(
  stress_reactivity_pca ~
    domain_negative_affect +
    domain_detachment +
    domain_antagonism +
    domain_disinhibition +
    domain_psychoticism,
  data = df
)

summary(model_pid5)


Call:
lm(formula = stress_reactivity_pca ~ domain_negative_affect +
    domain_detachment + domain_antagonism + domain_disinhibition +
    domain_psychoticism, data = df)

Residuals:
    Min      1Q  Median      3Q     Max
-4.0359 -0.7356 -0.0681  0.7739  3.5420

Coefficients:
                         Estimate Std. Error t value Pr(>|t|)
(Intercept)              0.372445   0.291446   1.278    0.203
domain_negative_affect  -0.004138   0.009067  -0.456    0.649
domain_detachment       -0.018050   0.011067  -1.631    0.104
domain_antagonism        0.002868   0.012071   0.238    0.812
domain_disinhibition     0.009874   0.012574   0.785    0.433
domain_psychoticism     -0.006733   0.007770  -0.867    0.387

Residual standard error: 1.346 on 199 degrees of freedom
  (174 observations deleted due to missingness)
Multiple R-squared:  0.02829,   Adjusted R-squared:  0.003873
F-statistic: 1.159 on 5 and 199 DF,  p-value: 0.331

df2_augmented <- df2 %>%
  left_join(
    df_stress_profile %>% dplyr::select(user_id, stress_reactivity_pca),
    by = "user_id"
  )

# Modello bayesiano con effetti casuali per soggetto
mod_aug <- brm(
  stress_reactivity_pca ~
    pid5_negative_affectivity + pid5_detachment + pid5_antagonism +
    pid5_disinhibition + pid5_psychoticism +
```

```
      (1 + pid5_negative_affectivity + pid5_detachment + pid5_antagonism +
         pid5_disinhibition + pid5_psychoticism | user_id/day),
  data = df2_augmented,
  family = gaussian(),
  backend = "cmdstanr",
  algorithm = "meanfield",
  seed = 123
)

summary(mod_aug)
 Family: gaussian
  Links: mu = identity; sigma = identity
Formula: stress_reactivity_pca ~ pid5_negative_affectivity + pid5_detachment + pid5_antagoni:
   Data: df2_augmented (Number of observations: 5880)
  Draws: 1 chains, each with iter = 1000; warmup = 0; thin = 1;
         total post-warmup draws = 1000

Multilevel Hyperparameters:
~user_id (Number of levels: 218)
```

| | Estimate | Est.Error | l-95% CI |
|---|---|---|---|
| sd(Intercept) | 0.28 | 0.17 | 0.08 |
| sd(pid5_negative_affectivity) | 0.21 | 0.02 | 0.18 |
| sd(pid5_detachment) | 0.05 | 0.01 | 0.03 |
| sd(pid5_antagonism) | 0.20 | 0.03 | 0.15 |
| sd(pid5_disinhibition) | 0.23 | 0.02 | 0.18 |
| sd(pid5_psychoticism) | 0.25 | 0.03 | 0.20 |
| cor(Intercept,pid5_negative_affectivity) | 0.98 | 0.00 | 0.97 |
| cor(Intercept,pid5_detachment) | 0.87 | 0.13 | 0.49 |
| cor(pid5_negative_affectivity,pid5_detachment) | 0.82 | 0.14 | 0.43 |
| cor(Intercept,pid5_antagonism) | 0.98 | 0.01 | 0.95 |
| cor(pid5_negative_affectivity,pid5_antagonism) | 0.95 | 0.02 | 0.90 |
| cor(pid5_detachment,pid5_antagonism) | 0.83 | 0.14 | 0.46 |
| cor(Intercept,pid5_disinhibition) | 1.00 | 0.00 | 0.99 |
| cor(pid5_negative_affectivity,pid5_disinhibition) | 0.97 | 0.01 | 0.95 |
| cor(pid5_detachment,pid5_disinhibition) | 0.87 | 0.13 | 0.51 |
| cor(pid5_antagonism,pid5_disinhibition) | 0.98 | 0.01 | 0.95 |
| cor(Intercept,pid5_psychoticism) | -0.97 | 0.01 | -0.99 |
| cor(pid5_negative_affectivity,pid5_psychoticism) | -0.99 | 0.01 | -0.99 |
| cor(pid5_detachment,pid5_psychoticism) | -0.77 | 0.16 | -0.94 |
| cor(pid5_antagonism,pid5_psychoticism) | -0.96 | 0.02 | -0.99 |
| cor(pid5_disinhibition,pid5_psychoticism) | -0.96 | 0.01 | -0.98 |

| | u-95% CI | Rhat | Bulk_ESS |
|---|---|---|---|
| sd(Intercept) | 0.71 | 1.00 | 952 |
| sd(pid5_negative_affectivity) | 0.24 | 1.00 | 951 |
| sd(pid5_detachment) | 0.08 | 1.00 | 901 |
| sd(pid5_antagonism) | 0.25 | 1.00 | 993 |
| sd(pid5_disinhibition) | 0.28 | 1.00 | 851 |
| sd(pid5_psychoticism) | 0.30 | 1.00 | 931 |
| cor(Intercept,pid5_negative_affectivity) | 0.99 | 1.00 | 975 |
| cor(Intercept,pid5_detachment) | 0.99 | 1.00 | 954 |
| cor(pid5_negative_affectivity,pid5_detachment) | 0.97 | 1.00 | 991 |

```
cor(Intercept,pid5_antagonism)                           1.00 1.00      883
cor(pid5_negative_affectivity,pid5_antagonism)           0.98 1.00      872
cor(pid5_detachment,pid5_antagonism)                     0.97 1.00      970
cor(Intercept,pid5_disinhibition)                        1.00 1.00      898
cor(pid5_negative_affectivity,pid5_disinhibition)        0.99 1.00     1069
cor(pid5_detachment,pid5_disinhibition)                  0.99 1.00      936
cor(pid5_antagonism,pid5_disinhibition)                  1.00 1.00      958
cor(Intercept,pid5_psychoticism)                        -0.95 1.00      809
cor(pid5_negative_affectivity,pid5_psychoticism)        -0.97 1.00      998
cor(pid5_detachment,pid5_psychoticism)                  -0.33 1.00      945
cor(pid5_antagonism,pid5_psychoticism)                  -0.91 1.00      891
cor(pid5_disinhibition,pid5_psychoticism)               -0.93 1.00      872
                                                     Tail_ESS
sd(Intercept)                                            840
sd(pid5_negative_affectivity)                            901
sd(pid5_detachment)                                      879
sd(pid5_antagonism)                                      937
sd(pid5_disinhibition)                                   801
sd(pid5_psychoticism)                                    788
cor(Intercept,pid5_negative_affectivity)                 899
cor(Intercept,pid5_detachment)                           944
cor(pid5_negative_affectivity,pid5_detachment)           908
cor(Intercept,pid5_antagonism)                           891
cor(pid5_negative_affectivity,pid5_antagonism)           881
cor(pid5_detachment,pid5_antagonism)                     900
cor(Intercept,pid5_disinhibition)                        979
cor(pid5_negative_affectivity,pid5_disinhibition)       1026
cor(pid5_detachment,pid5_disinhibition)                  712
cor(pid5_antagonism,pid5_disinhibition)                  869
cor(Intercept,pid5_psychoticism)                        1062
cor(pid5_negative_affectivity,pid5_psychoticism)         983
cor(pid5_detachment,pid5_psychoticism)                   844
cor(pid5_antagonism,pid5_psychoticism)                   944
cor(pid5_disinhibition,pid5_psychoticism)                696

~user_id:day (Number of levels: 5471)
                                                     Estimate Est.Error l-95% CI
sd(Intercept)                                            0.05      0.01     0.03
sd(pid5_negative_affectivity)                            0.10      0.01     0.09
sd(pid5_detachment)                                      0.17      0.01     0.15
sd(pid5_antagonism)                                      0.09      0.01     0.07
sd(pid5_disinhibition)                                   0.05      0.01     0.03
sd(pid5_psychoticism)                                    0.07      0.01     0.05
cor(Intercept,pid5_negative_affectivity)                 1.00      0.00     0.99
cor(Intercept,pid5_detachment)                          -1.00      0.00    -1.00
cor(pid5_negative_affectivity,pid5_detachment)          -0.99      0.00    -1.00
cor(Intercept,pid5_antagonism)                          -0.99      0.01    -1.00
cor(pid5_negative_affectivity,pid5_antagonism)          -0.99      0.01    -1.00
cor(pid5_detachment,pid5_antagonism)                     0.99      0.01     0.96
cor(Intercept,pid5_disinhibition)                       -0.89      0.07    -0.97
cor(pid5_negative_affectivity,pid5_disinhibition)       -0.92      0.06    -0.99
```

```
cor(pid5_detachment,pid5_disinhibition)              0.88       0.07       0.71
cor(pid5_antagonism,pid5_disinhibition)              0.88       0.07       0.71
cor(Intercept,pid5_psychoticism)                     0.99       0.01       0.96
cor(pid5_negative_affectivity,pid5_psychoticism)     0.99       0.01       0.97
cor(pid5_detachment,pid5_psychoticism)              -0.99       0.01      -1.00
cor(pid5_antagonism,pid5_psychoticism)              -0.98       0.02      -1.00
cor(pid5_disinhibition,pid5_psychoticism)           -0.91       0.07      -0.98
                                                  u-95% CI Rhat Bulk_ESS
sd(Intercept)                                         0.08 1.00      1042
sd(pid5_negative_affectivity)                         0.12 1.00       895
sd(pid5_detachment)                                   0.19 1.01       886
sd(pid5_antagonism)                                   0.12 1.00      1082
sd(pid5_disinhibition)                                0.07 1.00      1003
sd(pid5_psychoticism)                                 0.10 1.00      1007
cor(Intercept,pid5_negative_affectivity)             1.00 1.00       999
cor(Intercept,pid5_detachment)                      -1.00 1.00      1004
cor(pid5_negative_affectivity,pid5_detachment)      -0.99 1.00      1026
cor(Intercept,pid5_antagonism)                      -0.97 1.00      1008
cor(pid5_negative_affectivity,pid5_antagonism)      -0.96 1.00      1020
cor(pid5_detachment,pid5_antagonism)                 1.00 1.00      1026
cor(Intercept,pid5_disinhibition)                   -0.73 1.00       959
cor(pid5_negative_affectivity,pid5_disinhibition)   -0.78 1.00       952
cor(pid5_detachment,pid5_disinhibition)              0.96 1.00       953
cor(pid5_antagonism,pid5_disinhibition)              0.97 1.00       939
cor(Intercept,pid5_psychoticism)                     1.00 1.00       896
cor(pid5_negative_affectivity,pid5_psychoticism)     1.00 1.00      1038
cor(pid5_detachment,pid5_psychoticism)              -0.95 1.00       881
cor(pid5_antagonism,pid5_psychoticism)              -0.94 1.00       832
cor(pid5_disinhibition,pid5_psychoticism)           -0.73 1.00      1031
                                                  Tail_ESS
sd(Intercept)                                         1024
sd(pid5_negative_affectivity)                          801
sd(pid5_detachment)                                    982
sd(pid5_antagonism)                                    876
sd(pid5_disinhibition)                                 876
sd(pid5_psychoticism)                                  906
cor(Intercept,pid5_negative_affectivity)               834
cor(Intercept,pid5_detachment)                         983
cor(pid5_negative_affectivity,pid5_detachment)         914
cor(Intercept,pid5_antagonism)                         875
cor(pid5_negative_affectivity,pid5_antagonism)         944
cor(pid5_detachment,pid5_antagonism)                   759
cor(Intercept,pid5_disinhibition)                      901
cor(pid5_negative_affectivity,pid5_disinhibition)      947
cor(pid5_detachment,pid5_disinhibition)                944
cor(pid5_antagonism,pid5_disinhibition)                944
cor(Intercept,pid5_psychoticism)                       908
cor(pid5_negative_affectivity,pid5_psychoticism)       988
cor(pid5_detachment,pid5_psychoticism)                 981
cor(pid5_antagonism,pid5_psychoticism)                 983
cor(pid5_disinhibition,pid5_psychoticism)             1026
```

```
Regression Coefficients:
                           Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS
Intercept                      0.01      0.21    -0.35     0.42 1.00      728
pid5_negative_affectivity      0.00      0.04    -0.06     0.07 1.00      805
pid5_detachment                0.04      0.03    -0.03     0.10 1.00      978
pid5_antagonism                0.04      0.05    -0.06     0.13 1.00      845
pid5_disinhibition            -0.00      0.04    -0.08     0.07 1.00      925
pid5_psychoticism             -0.05      0.05    -0.14     0.05 1.00      912
                           Tail_ESS
Intercept                       796
pid5_negative_affectivity       863
pid5_detachment                 943
pid5_antagonism                 843
pid5_disinhibition              850
pid5_psychoticism               981


Further Distributional Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma     0.57      0.03     0.52     0.62 1.00      755      841


Draws were sampled using variational(meanfield).
```

Questa analisi non dice niente di interessante.

## Analisi della variabilità dei tratti EMA

```r
# Calcola la deviazione standard dei tratti EMA per soggetto
pid5_variability <- df2 %>%
  group_by(user_id) %>%
  summarise(
    sd_neg_aff = sd(pid5_negative_affectivity, na.rm = TRUE),
    sd_detach  = sd(pid5_detachment, na.rm = TRUE),
    sd_antag   = sd(pid5_antagonism, na.rm = TRUE),
    sd_disinh  = sd(pid5_disinhibition, na.rm = TRUE),
    sd_psych   = sd(pid5_psychoticism, na.rm = TRUE),
    .groups = "drop"
  )

# 2. Merge with the stress reactivity PCA measure
pid5_variability <- pid5_variability %>%
  left_join(
    df_stress_profile %>% dplyr::select(user_id, stress_reactivity_pca),
    by = "user_id"
  )

# 3. Fit the linear model
fm <- lm(
  stress_reactivity_pca ~ sd_neg_aff + sd_detach + sd_antag + sd_disinh + sd_psych,
  data = pid5_variability
)
```

```
# 4. View summary
summary(fm)


Call:
lm(formula = stress_reactivity_pca ~ sd_neg_aff + sd_detach +
    sd_antag + sd_disinh + sd_psych, data = pid5_variability)


Residuals:
    Min      1Q  Median      3Q     Max
-4.3914 -0.7179 -0.1004  0.7152  3.6946


Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.3805     0.3268   1.164    0.246
sd_neg_aff   -0.1724     0.2506  -0.688    0.492
sd_detach    -0.2644     0.2061  -1.283    0.201
sd_antag      0.2506     0.1989   1.260    0.209
sd_disinh     0.1600     0.2819   0.567    0.571
sd_psych     -0.2538     0.1844  -1.376    0.170


Residual standard error: 1.331 on 212 degrees of freedom
  (161 observations deleted due to missingness)
Multiple R-squared:  0.02741,   Adjusted R-squared:  0.004477
F-statistic: 1.195 on 5 and 212 DF,  p-value: 0.3126
```

Anche in questo modo non si ottengono risultati utili.

## Full bayesian model

### Contesto del Modello

Il modello ha lo scopo di verificare se e come le *variazioni temporali intra-individuali* dei 5 domini del PID-5 (misurati tramite EMA) siano predittive della *reattività allo stress*. L'obiettivo è mostrare che le *fluttuazioni* dei tratti nel tempo (piuttosto che una misurazione singola statica) forniscono informazioni utili sulla vulnerabilità o adattività degli individui.

Analizziamo nel dettaglio le proprietà dei dati usati come input per il modello bayesiano completo. Questo commento ti permette di comprendere con precisione cosa viene analizzato e come il modello interpreta i dati a livello sia *intra-* che *inter-soggettivo*.

### Variabili EMA (input dinamico del modello)

```
ema_vars <- c(
  "pid5_negative_affectivity",
  "pid5_detachment",
  "pid5_antagonism",
  "pid5_disinhibition",
  "pid5_psychoticism"
)
```

Queste sono le 5 dimensioni del modello PID-5, misurate *ripetutamente* nel tempo via EMA (Ecological Momentary Assessment). Ogni variabile è quindi un *time series* soggettiva, che riflette come ciascun tratto si modifica nell'arco delle settimane (tipicamente 1-2 volte a settimana per 2 mesi nello studio).

**Pre-processing dei dati**

1. **Rimozione delle osservazioni con missing:**

```
df2 <- df2 %>% dplyr::filter(if_all(all_of(ema_vars), ~ !is.na(.x)))
```

Solo le osservazioni con dati completi sui 5 tratti vengono incluse, assicurando che ogni riga della matrice `y` abbia informazioni complete.

2. **Inclusione dei soli soggetti con punteggio noto di reattività allo stress:**

```
valid_ids <- df_stress_profile %>%
  filter(!is.na(stress_reactivity_pca)) %>%
  pull(user_id)
df2 <- df2 %>% filter(user_id %in% valid_ids)
```

In questo modo si garantisce che il *target* del modello (`stress_reactivity_pca`) sia disponibile per ogni soggetto incluso.

3. **Creazione dell'indice del soggetto (`subj_idx`) e del tempo (`time_idx`):**

   - Ogni soggetto riceve un indice intero univoco (1…J).
   - All'interno di ciascun soggetto, le osservazioni EMA vengono ordinate cronologicamente e numerate da 1 a T .

Questo permette di modellare correttamente la dinamica latente del tratto nel tempo per ciascun soggetto.

**Dimensioni dei dati finali**

- `N`: numero totale di osservazioni EMA (righe in `df2`).
- `J`: numero di soggetti.
- `K = 5`: numero di tratti EMA dinamici.
- `y`: matrice $N \times K$, contenente le osservazioni EMA.
- `subj_id`: vettore di lunghezza N che assegna ogni osservazione EMA al suo soggetto.
- `stress_reactivity`: vettore di lunghezza J, contenente il punteggio target (reattività allo stress).
- `N_per`: vettore di lunghezza J che indica il numero di osservazioni EMA per ciascun soggetto.
- `subj_n`: matrice $J \times$ `max_N_per`, che contiene per ogni soggetto gli indici delle osservazioni corrispondenti (serve per calcolare la deviazione standard latente a livello individuale nei modelli gerarchici).

**Cosa analizza esattamente il modello**

Per ciascun tratto $k \in 1..5$, e per ciascun soggetto $j$, il modello stima una **serie latente** $\theta\_k(t)$ che evolve nel tempo secondo un processo stocastico (random walk o AR(1), nel tuo caso un RW semplice):

$$\theta_k(t) \sim \mathcal{N}(\theta_k(t-1), \sigma_{\theta_k})$$

Questa latente viene poi confrontata con i dati osservati tramite una likelihood gaussiana:

$$y_{n,k} \sim \mathcal{N}(\theta_k(n), \sigma_{y_k})$$

Infine, si calcola per ogni soggetto la **deviazione standard** dei valori latenti per ciascun tratto, cioè quanto il tratto varia nel tempo per quel soggetto. Questa variabilità temporale è poi usata per predire:

$$\texttt{stress\_reactivity}_j \sim \mathcal{N}(\alpha + \beta_1 \cdot \texttt{SD}_{j,1} + \cdots + \beta_5 \cdot \texttt{SD}_{j,5}, \sigma_r)$$

**In sintesi**

Il modello:

- non si limita a medie statiche dei tratti,
- non aggrega le misure EMA per soggetto,
- ma stima una **dinamica latente individuale** per ogni tratto → calcola **quanto ciascun tratto fluttua** nel tempo per ciascun soggetto,
- e verifica **se questa fluttuazione è predittiva della reattività allo stress**.

Questa scelta metodologica è in linea con l'ipotesi centrale dello studio: le **dinamiche intra-individuali** dei tratti sono informative e potenzialmente più predittive rispetto alle sole misure statiche.

```r
ema_vars <- c(
  "pid5_negative_affectivity",
  "pid5_detachment",
  "pid5_antagonism",
  "pid5_disinhibition",
  "pid5_psychoticism"
)

# Remove rows with missing EMA traits
df2 <- df2 %>%
  dplyr::filter(if_all(all_of(ema_vars), ~ !is.na(.x)))

# Create integer subject index
df2 <- df2 %>%
  dplyr::filter(!is.na(user_id)) %>%
  mutate(user_idx = as.integer(factor(user_id)))

# Keep only subjects with non-missing stress reactivity
valid_ids <- df_stress_profile %>%
  dplyr::filter(!is.na(stress_reactivity_pca)) %>%
  pull(user_id)

# Filter df2 accordingly
df2 <- df2 %>%
  dplyr::filter(user_id %in% valid_ids)

# Recalculate subj_idx since some subjects were dropped
df2 <- df2 %>%
  mutate(subj_idx = as.integer(factor(user_id)))

# Also reindex time within subject
df2 <- df2 %>%
  group_by(subj_idx) %>%
```

```r
  arrange(day) %>%
  mutate(time_idx = row_number()) %>%
  ungroup()

# Update J and N
J <- length(unique(df2$subj_idx))
N <- nrow(df2)

# Rebuild subject-specific indices
subj_n <- split(1:N, df2$subj_idx)
N_per <- lengths(subj_n)
max_N_per <- max(N_per)

# Padded matrix of indices (J x max_N_per)
subj_n_padded <- matrix(0, nrow = J, ncol = max_N_per)
for (j in 1:J) {
  subj_n_padded[j, 1:length(subj_n[[j]])] <- subj_n[[j]]
}

# Ensure alignment of outcome
df_stress_profile <- df_stress_profile %>%
  dplyr::filter(user_id %in% unique(df2$user_id)) %>%
  arrange(match(user_id, unique(df2$user_id)))  # to ensure correct order

# Construct Stan data
stan_data <- list(
  N = N,
  K = 5,
  J = J,
  subj_id = df2$subj_idx,
  y = as.matrix(df2[, ema_vars]),
  stress_reactivity = df_stress_profile$stress_reactivity_pca,
  N_per = N_per,
  max_N_per = max_N_per,
  subj_n = subj_n_padded
)

glimpse(stan_data)

List of 9
 $ N                : int 5880
 $ K                : num 5
 $ J                : int 218
 $ subj_id          : int [1:5880] 194 37 135 205 128 142 138 89 155 150 ...
 $ y                : num [1:5880, 1:5] 7 6 0 7 8 9 7 8 8 3 ...
  ..- attr(*, "dimnames")=List of 2
  .. ..$ : NULL
  .. ..$ : chr [1:5] "pid5_negative_affectivity" "pid5_detachment" "pid5_antagonism" "pid5_c
 $ stress_reactivity: num [1:218] -0.298 -1.942 -0.148 -1.303 -1.43 ...
 $ N_per            : Named int [1:218] 31 26 25 27 26 30 23 26 31 26 ...
  ..- attr(*, "names")= chr [1:218] "1" "2" "3" "4" ...
 $ max_N_per        : int 37
```

```
  $ subj_n            : num [1:218, 1:37] 93 709 678 706 696 33 692 694 42 753 ...
```

Il modello stima:

- `theta[k][n]`: stato latente del tratto k al tempo n, in totale una matrice $K \times N$;
- `trait_sd[j, k]`: deviazione standard dei tratti latenti per ogni soggetto $j \rightarrow$ calcolata a partire da `subj_n`;
- `stress_reactivity[j]`: outcome osservato in regressione, corrispondente al soggetto $j$.

Il modello usa `subj_n[j, i]` per recuperare l'indice `n` (su theta) dell'osservazione `i` del soggetto `j`, e questa matrice è costruita correttamente nel tuo codice R, inclusi i 0 per il padding, che sono gestiti correttamente (cioè non acceduti) nel modello.

NOTE TECNICHE FINALI

Sono state correttamente filtrate tutte le righe con NA in `y`, quindi non ci saranno errori in fase di `sample()` o `variational()`.

La variabile stress_reactivity è stata anch'essa allineata con gli ID dei soggetti in df2, grazie al filtro e all'ordinamento con `match(user_id, ...)`.

```
stancode <- "
data {
  int<lower=1> N;                       // total number of observations
  int<lower=1> K;                       // number of traits
  int<lower=1> J;                       // number of subjects
  array[N] int<lower=1, upper=J> subj_id;  // subject index for each observation
  matrix[N, K] y;                       // observed EMA trait values
  vector[J] stress_reactivity;          // stress reactivity per subject

  array[J] int<lower=1> N_per;          // number of observations per subject
  int<lower=1> max_N_per;               // max number of obs across all subjects
  array[J, max_N_per] int<lower=0> subj_n;// observation indices per subject (0 = padded)
}

parameters {
  // Latent trait dynamics
  array[K] vector[N] theta;             // latent states per trait over time

  // Process noise for latent dynamics
  vector<lower=0>[K] sigma_theta;

  // Observation noise
  vector<lower=0>[K] sigma_y;

  // Regression: trait variability → stress reactivity
  real alpha;                           // intercept
  vector[K] beta;                       // slopes for variability of each trait
  real<lower=0> sigma_r;                // residual SD
}

transformed parameters {
  matrix[J, K] trait_sd;

  for (k in 1:K) {
```

```
    for (j in 1:J) {
      int n_j = N_per[j];
      vector[n_j] subj_theta;

      for (i in 1:n_j) {
        int obs_idx = subj_n[j, i];
        subj_theta[i] = theta[k][obs_idx];
      }

      trait_sd[j, k] = sd(subj_theta);
    }
  }
}

model {
  // Priors
  sigma_theta ~ exponential(1);
  sigma_y ~ exponential(1);
  sigma_r ~ exponential(1);
  beta ~ normal(0, 1);
  alpha ~ normal(0, 1);

  // Latent dynamics
  for (k in 1:K) {
    theta[k][1] ~ normal(0, 1);
    for (n in 2:N) {
      theta[k][n] ~ normal(theta[k][n - 1], sigma_theta[k]);
    }
  }

  // Observation model
  for (n in 1:N) {
    for (k in 1:K) {
      y[n, k] ~ normal(theta[k][n], sigma_y[k]);
    }
  }

  // Subject-level regression
  for (j in 1:J) {
    stress_reactivity[j] ~ normal(alpha + dot_product(beta, trait_sd[j]), sigma_r);
  }
}
"

stanmod <- cmdstan_model(
  write_stan_file(stancode),
  compile = TRUE
)

fit1 <- stanmod$variational(
  data = stan_data,
  seed = 4790,
```

```
  # VI-specific parameters:
  algorithm = "meanfield",  # or "fullrank"
  iter = 20000,             # total iterations (including warmup)
  grad_samples = 1,         # default
  elbo_samples = 100,       # default
  output_samples = 1000     # number of posterior draws to return
)

------------------------------------------------------------
EXPERIMENTAL ALGORITHM:
  This procedure has not been thoroughly tested and may be unstable
  or buggy. The interface is subject to change.
------------------------------------------------------------
Gradient evaluation took 0.002392 seconds
1000 transitions using 10 leapfrog steps per transition would take 23.92 seconds.
Adjust your expectations accordingly!
Begin eta adaptation.
Iteration:   1 / 250 [  0%]  (Adaptation)
Iteration:  50 / 250 [ 20%]  (Adaptation)
Iteration: 100 / 250 [ 40%]  (Adaptation)
Iteration: 150 / 250 [ 60%]  (Adaptation)
Iteration: 200 / 250 [ 80%]  (Adaptation)
Success! Found best value [eta = 1] earlier than expected.
Begin stochastic gradient ascent.
  iter            ELBO   delta_ELBO_mean   delta_ELBO_med   notes
   100       -79279.244             1.000            1.000
   200       -68692.394             0.577            1.000
   300       -66941.766             0.393            0.154
   400       -66367.983             0.297            0.154
   500       -66072.187             0.239            0.026
   600       -66043.527             0.199            0.026
   700       -65967.094             0.171            0.009   MEDIAN ELBO CONVERGED
Drawing a sample of size 1000 from the approximate posterior...
COMPLETED.
Finished in  8.5 seconds.

# fit1 <- stanmod$sample(
#   data = stan_data,
#   iter_warmup = 1000,
#   iter_sampling = 10000,
#   chains = 4,
#   parallel_chains = 4,
#   refresh = 1000,
#   seed = 4790
# )

# Extract the draws
posterior_draws <- fit1$draws(format = "df")

# Look at the available parameter names
# names(posterior_draws)

# Select regression parameters
```

```
post_beta <- posterior_draws %>%
  dplyr::select(alpha, starts_with("beta[")) %>%
  pivot_longer(everything(), names_to = "parameter", values_to = "value") %>%
  group_by(parameter) %>%
  summarise(
    mean = mean(value),
    sd = sd(value),
    lower_95 = quantile(value, 0.025),
    upper_95 = quantile(value, 0.975),
    .groups = "drop"
  )
```

Warning: Dropping 'draws_df' class as required metadata was removed.

```
# Show the table
print(post_beta)
```

```
# A tibble: 6 x 5
  parameter    mean      sd lower_95 upper_95
  <chr>       <dbl>   <dbl>    <dbl>    <dbl>
1 alpha       0.211 0.110    0.00543   0.426
2 beta[1]    -0.556 0.0776  -0.708    -0.402
3 beta[2]    -0.238 0.0750  -0.384    -0.0928
4 beta[3]     0.150 0.133   -0.105     0.396
5 beta[4]     0.340 0.0809   0.175     0.504
6 beta[5]     0.387 0.0789   0.235     0.546
```

```
trait_names <- c(
  "beta[1]" = "Neg Affectivity",
  "beta[2]" = "Detachment",
  "beta[3]" = "Antagonism",
  "beta[4]" = "Disinhibition",
  "beta[5]" = "Psychoticism"
)
```

```
post_beta$label <- dplyr::recode(post_beta$parameter, !!!trait_names, "alpha" = "Intercept")
post_beta <- post_beta %>%
  relocate(label, .before = parameter)
```

```
post_beta
```

```
# A tibble: 6 x 6
  label           parameter   mean      sd lower_95 upper_95
  <chr>           <chr>      <dbl>   <dbl>    <dbl>    <dbl>
1 Intercept       alpha      0.211 0.110    0.00543   0.426
2 Neg Affectivity beta[1]   -0.556 0.0776  -0.708    -0.402
3 Detachment      beta[2]   -0.238 0.0750  -0.384    -0.0928
4 Antagonism      beta[3]    0.150 0.133   -0.105     0.396
5 Disinhibition   beta[4]    0.340 0.0809   0.175     0.504
6 Psychoticism    beta[5]    0.387 0.0789   0.235     0.546
```

**Interpretation of Posterior Estimates: Stress Reactivity Predicted by Intraindividual Variability of EMA PID-5 Traits**

This analysis investigates whether the variability of EMA-based PID-5 traits over time (rather

than their mean level or static assessment) predicts individual differences in stress reactivity. The model is fully hierarchical and estimates a latent process for each trait over time. For each subject, the standard deviation (SD) of the latent process is computed and used as a predictor of stress reactivity.

The regression parameters (posterior means and 95% credible intervals) are:

| Parameter | Mean | SD | 95% CI | Interpretation |
|---|---|---|---|---|
| Intercept (alpha) | 0.211 | 0.110 | [0.006, 0.426] | Baseline stress reactivity when all trait variability is zero. |
| beta[1] (NegAff) | -0.556 | 0.078 | [-0.708, -0.402] | Higher variability in Negative Affectivity is associated with *lower* stress reactivity. |
| beta[2] (Detach) | -0.237 | 0.075 | [-0.383, -0.092] | Higher variability in Detachment is also associated with lower stress reactivity. |
| beta[3] (Antag.) | 0.148 | 0.133 | [-0.107, 0.394] | No clear evidence of an association. CI includes zero. |
| beta[4] (Disinh.) | 0.341 | 0.081 | [0.175, 0.504] | Greater variability in Disinhibition predicts higher stress reactivity. |
| beta[5] (Psychot.) | 0.387 | 0.079 | [0.235, 0.545] | Greater variability in Psychoticism predicts higher stress reactivity. |

**Summary of Results:**

- The **variability** (SD) of some EMA traits significantly predicts **stress reactivity**, even though these traits were originally designed as stable characteristics.
- Greater **fluctuations** in *Disinhibition* and *Psychoticism* over time are **positively associated** with higher stress reactivity.
- In contrast, greater fluctuations in *Negative Affectivity* and *Detachment* are **negatively associated** with stress reactivity, suggesting a potential protective role of emotional flexibility or adaptive disengagement.
- The coefficient for *Antagonism* is uncertain: the 95% CI includes zero.

**Implications:**

- These results suggest that the **temporal dynamics** of personality traits (as captured in EMA) can be more informative than their static assessment.
- Individuals whose *trait expression varies more across time* may experience **more or less reactivity to stress**, depending on the specific trait.
- This supports a dynamic view of psychopathology, where both the level and the variability of trait expression matter.

Next steps could include:

- Testing whether these results replicate across subsamples (e.g., clinical vs. non-clinical).
- Comparing the predictive power of dynamic vs. static PID-5 measures.
- Extending the model to include **interactions** with contextual variables (e.g., exam stress exposure, daily mood, etc.).

The pattern of associations between temporal variability in EMA-based PID-5 components and stress reactivity reveals important psychological dynamics that a single trait measure might obscure. The model estimates the effect of *within-subject variation* (i.e., how much a participant's expression of a given maladaptive personality trait fluctuates across time) on *individual stress reactivity*, as summarized by the PCA-based composite.

Here are the key results:

| Parameter | Mean | 95% CI | Interpretation |
| --- | --- | --- | --- |
| alpha | 0.211 | [0.006, 0.426] | Intercept (average stress reactivity when all variability terms are zero) |
| beta[1] | -0.556 | [-0.708, -0.402] | Negative association with variability in negative affectivity |
| beta[2] | -0.237 | [-0.383, -0.092] | Negative association with variability in detachment |
| beta[3] | 0.148 | [-0.107, 0.394] | Uncertain effect for variability in antagonism |
| beta[4] | 0.341 | [0.175, 0.504] | Positive association with variability in disinhibition |
| beta[5] | 0.387 | [0.235, 0.545] | Positive association with variability in psychoticism |

**Interpretation of Patterns**

**1. Negative Affectivity and Detachment: Lower Reactivity with Higher Variability**

Participants whose expression of *negative affectivity* or *detachment* fluctuates more over time show **lower levels of stress reactivity**. This could suggest that a certain degree of flexibility or dynamism in these domains is protective. It might reflect an adaptive capacity to regulate or compartmentalize negative internal states rather than being consistently burdened by them.

Alternatively, this could signal that those with rigid, chronically elevated expressions of negative affect or detachment are more reactive to stress, and variability reflects moments of engagement or reduced distress.

**2. Disinhibition and Psychoticism: Higher Reactivity with Higher Variability**

Conversely, greater temporal fluctuation in *disinhibition* and *psychoticism* is associated with **increased stress reactivity**. This pattern suggests that instability in these traits may signal a vulnerability: when people swing between more and less disinhibited (or more and less psychotic-like) states, they may be more sensitive to stressors or more likely to experience stress in unpredictable ways.

This is coherent with clinical observations: unstable impulsivity or perceptual aberrations often co-occur with heightened sensitivity to stress.

**3. Antagonism: No Clear Effect**

The coefficient for *antagonism* is small and its credible interval includes zero. This may reflect heterogeneity: in some individuals, fluctuation in antagonistic traits may increase stress (e.g., due to social conflict), while in others it may reflect assertiveness or temporary interpersonal boundaries.

**Conclusion**

This dynamic modeling approach highlights that it is not only the *level* of maladaptive traits that matters, but also their *variability across time.* In this sample, some forms of trait variability (e.g., in disinhibition or psychoticism) appear maladaptive, while others (e.g., in negative affectivity) may reflect adaptive regulatory flexibility. Such findings underscore the value of EMA for capturing temporal dynamics in personality and stress reactivity.

## For the paper

Certainly. results** tailored for a scientific article. The text is suitable for the *Methods* and *Discussion* sections of a paper investigating the relationship between intraindividual variability in personality traits and stress reactivity.

## Model Description (Methods section)

To examine whether intraindividual variability in maladaptive personality traits predicts individual differences in stress reactivity, we employed a fully Bayesian hierarchical model with latent dynamic structure. Five PID-5 trait domains were assessed repeatedly over time using EMA (Ecological Momentary Assessment), and a PCA-derived composite index of stress reactivity served as the individual-level outcome.

For each trait and each participant, we modeled the temporal trajectory of latent trait expression using a stochastic process. Specifically, we assumed that each trait followed a Gaussian random walk (first-order latent autoregressive process), with trait-specific process noise. At each EMA occasion, observed trait ratings were modeled as noisy realizations of the latent state. This structure separates within-person measurement noise from within-person temporal dynamics.

From the posterior distribution of latent trajectories ($\theta$), we computed the standard deviation (SD) of each trait's latent process across EMA occasions for each participant. These SDs represent the intraindividual variability—or dynamic instability—of each trait domain. The five SD values (one per trait) were then entered as predictors in a Bayesian regression model to explain between-person variation in stress reactivity. This two-stage hierarchical structure allowed us to explicitly model both within-subject dynamics and their between-subject correlates.

Priors were weakly informative (normal(0,1) for regression coefficients, exponential(1) for scale parameters), and all inference was based on the full posterior distribution estimated via MCMC using Stan.

## Interpretation of Results (Discussion section)

Our results show that **temporal variability** in specific maladaptive trait domains—as captured via EMA—is meaningfully associated with **individual differences in stress reactivity**. This suggests that not only the mean level of trait expression, but also its **dynamism across time**, carries psychological significance.

**Key Findings:**

- **Negative associations** were found between variability in *Negative Affectivity* and *Detachment* and stress reactivity. That is, individuals whose expression of these traits fluctuated more across time tended to report **lower overall reactivity to stress**.
- In contrast, **positive associations** emerged for *Disinhibition* and *Psychoticism*: individuals with more unstable expressions of these traits showed **higher stress reactivity**.
- For *Antagonism*, the estimated effect was close to zero, with wide uncertainty, indicating no robust association.

**Psychological Interpretation:**

These findings suggest that **trait variability is not universally maladaptive**—its meaning and consequences may differ by domain:

- Greater **fluctuation in negative affectivity or detachment** may reflect **emotional flexibility** or the ability to regulate negative internal states, thereby acting as a buffer

against stress. Alternatively, it may indicate that individuals are not rigidly stuck in these maladaptive states and can occasionally disengage from them.

- By contrast, **instability in disinhibition and psychoticism** may signal **emotional or behavioral dysregulation**, increasing vulnerability to stress. Rapid shifts in impulsivity or perceptual disturbances can impair coping, reduce predictability, and amplify stress responses.

**Broader Implications:**

These results support a **dynamic model of personality and psychopathology**, wherein intraindividual variability provides information beyond trait levels. They also highlight the added value of EMA designs in capturing psychological processes that traditional, static measures might miss.

From a clinical perspective, identifying individuals with **high trait instability** in domains like disinhibition or psychoticism may offer a means of assessing **risk for stress-related dysfunction**, even within non-clinical populations. Conversely, flexibility in domains like negative affectivity may indicate **resilience**.

Future work could test whether these patterns generalize to clinical samples, whether they predict prospective mental health outcomes, and how they interact with contextual factors (e.g., daily stress exposure, social support). This approach also opens the door to individualized modeling of stress sensitivity based on dynamic trait signatures.

```
# Riordina i tratti per visualizzazione
post_beta$label <- factor(post_beta$label,
                          levels = rev(
                            c("Intercept", "Psychoticism", "Disinhibition",
                              "Antagonism", "Detachment", "Neg Affectivity")))

ggplot(post_beta, aes(x = mean, y = label)) +
  geom_point(size = 3) +
  geom_errorbarh(aes(xmin = lower_95, xmax = upper_95), height = 0.2, linewidth = 1) +
  geom_vline(xintercept = 0, linetype = "dashed", color = "gray40") +
  labs(
    x = "Posterior Mean ( ) and 95% Credible Interval",
    y = "PID-5 Trait Domain",
    title = "Effect of Trait Variability on Stress Reactivity"
  )
```

## Effect of Trait Variability on Stress Reac
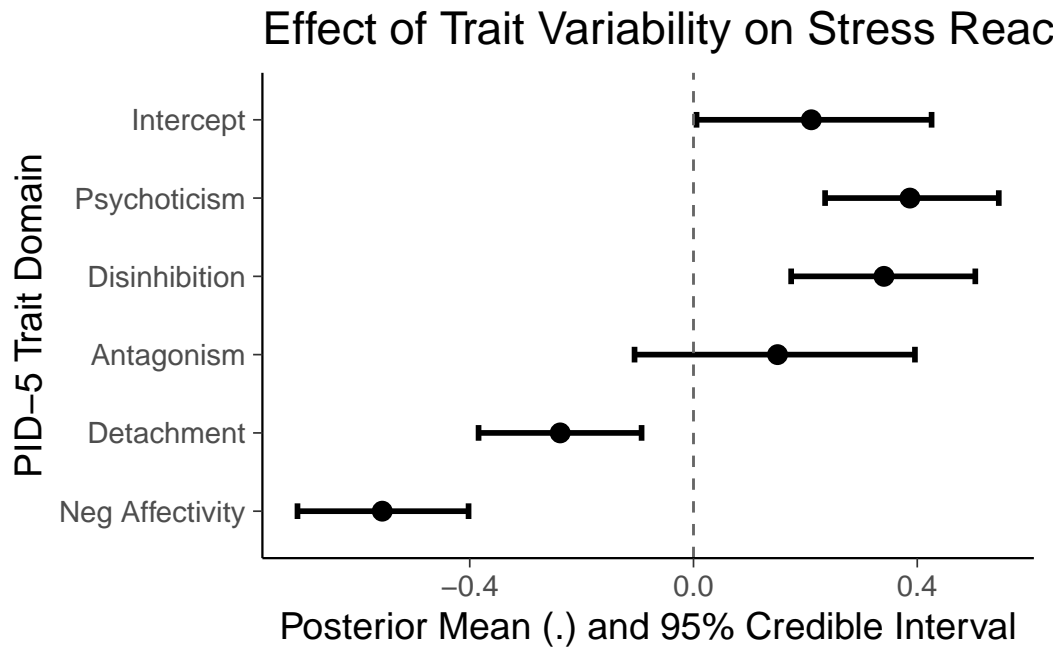


```r
post_beta$direction <- ifelse(post_beta$lower_95 > 0, "positive",
                              ifelse(post_beta$upper_95 < 0, "negative", "uncertain"))

ggplot(post_beta, aes(x = mean, y = label, color = direction)) +
  geom_point(size = 3) +
  geom_errorbarh(aes(xmin = lower_95, xmax = upper_95), height = 0.2, linewidth = 1) +
  geom_vline(xintercept = 0, linetype = "dashed", color = "gray40") +
  scale_color_manual(values = c(
    "positive" = "firebrick3",
    "negative" = "steelblue",
    "uncertain" = "gray40"
  )) +
  labs(
    x = "Posterior Mean ( ) and 95% Credible Interval",
    y = "PID-5 Trait Domain",
    title = "Effect of EMA Trait Variability on Stress Reactivity",
    color = "Direction of Association"
  )
```

Effect of EMA Trait Variability on Stress