

# Power Analysis for Multilevel Models With Cross-level Interactions in Intensive Longitudinal Designs

Ginette Lafit

2023-06-06

## Table of contents

Setting things up . . . . .	1
Description . . . . .	3
Determining model parameter values . . . . .	4
Analytical-based power analysis . . . . .	6
Step 1 . . . . .	6
Step 2 . . . . .	7
Step 3 . . . . .	10
Simulation-based power analysis . . . . .	10
Step 1 . . . . .	11
Step 2 . . . . .	11
Step 3 . . . . .	14
Session information . . . . .	14
References . . . . .	15

## Setting things up

Before we proceed, we need to ensure we have several packages installed and loaded into our R session. For the scripts below, we will use the following packages:

- [tidyverse](#)
- [gridExtra](#)
- [formattable](#)
- [htmltools](#)
- [shiny](#)

- DT
- ggplot2
- plyr
- dplyr
- tidyr
- shinyjs
- shinythemes
- viridis
- plotly
- remotes
- nlme
- devtools
- data.table
- MASS
- compiler
- future.apply

Which we can install in one go as follows:

```
# Prepare the package list.
packages = c(
  "tidyverse", "gridExtra", "formattable", "htmltools", "shiny",
  "DT", "ggplot2", "plyr", "dplyr", "tidyr", "shinyjs", "shinythemes",
  "viridis", "plotly", "remotes", "nlme", "devtools", "data.table",
  "MASS", "compiler", "future.apply"
)

# Install packages.
install.packages(packages)
```

#### Tip

You may consider first checking if the packages are installed before actually installing them. Nevertheless, the code above will not reinstall packages that are already installed and up-to-date.

At last, we can load the packages into our R session:

```
# Load packages.
library(tidyverse)
library(gridExtra)
library(formattable)
```

```

library(htmltools)
library(shiny)
library(DT)
library(ggplot2)
library(plyr)
library(dplyr)
library(tidyr)
library(shinyjs)
library(shinythemes)
library(viridis)
library(plotly)
library(remotes)
library(nlme)
library(devtools)
library(data.table)
library(MASS)
library(compiler)
library(future.apply)

```

## Description

The goal of this exercise is to conduct a power analysis to select the number of persons to investigate if depression moderates the relationship between anhedonia and negative affect (NA).

To estimate this research question we will use a multilevel model with a cross-level interaction effect between a level 1 continuous predictor (anhedonia) and level 2 continuous predictor (depression):

$$\text{NA}_{it} = \beta_{00} + \beta_{01}\text{Depression}_i + \beta_{10}\text{Anhedonia}_{it} + \beta_{01}\text{Depression}_i\text{Anhedonia}_{it} + \nu_{0i} + \nu_{1i}\text{Anhedonia}_{it} + \epsilon_{it} \quad (1)$$

where  $\beta_{00}$  is the fixed intercept,  $\beta_{01}$  represents the main effect of *depression*,  $\beta_{10}$  is the fixed slope of *anhedonia*, and  $\beta_{11}$  represents the cross-level interaction effect between *anhedonia* and *depression*. The cross-level interaction effect assesses whether *depression* moderates the effect of *anhedonia* on NA. The level 1 predictor (i.e., *anhedonia*) is centered within-persons and within-days.

In this model, we account for the serial dependency that characterizes IL designs by assuming that the level 1 errors follow an autoregressive AR(1) process (see Goldstein et al., 1994):

$$\epsilon_{it} = \rho_{\epsilon} \epsilon_{it-1} + \varepsilon_{it}$$

where the correlation between two consecutive errors is denoted by  $\rho_{\epsilon}$ , and  $\varepsilon_{it}$  is a Gaussian error with mean zero and variance  $\sigma_{\epsilon}^2$ . Under this model, the variance of the level 1 errors is given by  $\sigma_{\epsilon}^2/(1-\rho_{\epsilon}^2)$ . To guarantee that this model is stationary, the autocorrelation  $\rho_{\epsilon}$  should range between -1 and 1 (Hamilton, 1994).

Between-person differences in the relation between *anhedonia* and NA are captured by including a random intercept  $\nu_{0i}$  and random slope  $\nu_{1i}$ . These random effects are multivariate normal distributed with mean zero and covariance matrix  $\Sigma_{\nu}$ :

$$\Sigma_{\nu} = \begin{bmatrix} \sigma_{\nu_0}^2 & \sigma_{\nu_{01}} \\ \sigma_{\nu_{01}} & \sigma_{\nu_1}^2 \end{bmatrix}$$

In this model, it is also assumed that the level 2 random effects and the Level 1 errors are independent.

To investigate if depression moderates the relationship between anhedonia and NA we are going to conduct the following hypothesis test:

$$H_0 : \beta_{11} = 0$$

$$H_1 : \beta_{11} \neq 0$$

The aim of this exercise is to select the number of participants assuming the number of repeated measurements occasions to  $T = 70$ .

- Select sample size using the analytic approach, e.g.,  $N = \{20, 40, \dots\}$ .
- Compare the results with the ones obtained using the simulation-based approach.

## Determining model parameter values

To obtain the values of the model parameters we will use data from the Leuven clinical study. The code to estimate the multilevel model with cross-level interaction effect is included in the exercise [Multilevel model estimation using the Leuven Clinical Dataset](#). The output of the fitted model is:

Thus, the values of the model parameter that will be used to conduct the power analysis are:

```

Estimation output
## Random effects:
## Formula: ~1 + anhedonia.c | PID
## Structure: General positive-definite, Log-Cholesky parametrization
##           StdDev      Corr
## (Intercept) 12.855036 (Intr)
## anhedonia.c  0.1056154 0.249
## Residual    11.9234081
##
## Correlation Structure: AR(1)
## Formula: ~1 | PID
## Parameter estimate(s):
##      Phi
## 0.4302492
## Fixed effects: NA. ~ 1 + anhedonia.c + anhedonia.c * QIDS.c
##           Value Std.Error   DF   t-value p-value
## (Intercept)   42.97796 2.1228637 2215 20.245274  0.0000
## anhedonia.c     0.13747 0.0218391 2215  6.294553  0.0000
## QIDS.c          1.52600 0.4308459   36  3.541870  0.0011
## anhedonia.c:QIDS.c -0.01019 0.0046382 2215 -2.197910  0.0281
Mean anhedonia: 51.66162           Mean depression: 15.71
Std. deviation anhedonia: 23.6734   Std. deviation depression: 5.00

```

Figure 1: Ginette, can you please add the figure caption here?

$\beta_{00} = 42.98$  fixed intercept  
 $\beta_{01} = 1.53$  main effect of depression  
 $\beta_{10} = 0.14$  fixed slope  
 $\beta_{10} = -0.01$  cross-level interaction effect  
 $\sigma_{\epsilon} = 11.92$  std. deviation level 1 errors  
 $\rho_{\epsilon} = 0.43$  std. deviation level 1 errors  
 $\sigma_{\nu_0} = 12.86$  std. deviation random intercept  
 $\sigma_{\nu_1} = 0.11$  std. deviation random slope  
 $\rho_{\nu_{01}} = 0.249$  correlation between the random effects  
 $\mu_{\text{Anhedonia}} = 51.66$  mean anhedonia  
 $\sigma_{\text{Anhedonia}} = 23.67$  std. deviation anhedonia  
 $\mu_{\text{Depression}} = 15.71$  mean anhedonia  
 $\sigma_{\text{Depression}} = 5.00$  std. deviation anhedonia

## Analytical-based power analysis

To conduct the power analysis using the analytic approach we are going to use **ApproxPowerIL**: a **Shiny** application and R package to perform power analysis to select the number of persons for multilevel models with auto-correlated errors using asymptotic approximations of the information matrix. The repository contains functions used in Lafit et al. (2023). Users can download the app and run locally on their computer by executing the following commands in R or RStudio at [ApproxPowerIL](#).

```

# Install the app package from the `GitHub` repository.
remotes::install_github("ginettelafit/ApproxPowerIL", force = TRUE)

# Load package `ApproxPowerIL`.
library(ApproxPowerIL)

# Lunch the app from the `GitHub` gist.
shiny::runGist('302737dc046b89b7f09d15843389161c')

```

### Step 1

Select the model and set the sample size in the **ApproxPowerIL** application.

- Indicate the model of interest.

- Input the number of participants  $N$  (comma-separated):  $N = \{20, 40, 60, 80, 100\}$ .
- Input the number of repeated measurement occasions:  $T = 70$ .

### Choose multilevel model:

Model 5: Cross-level interaction effects (random slope) ▼

Model 5: Cross-level interaction effects (random slope)

Level 1:  $Y_{it} = \gamma_{0i} + \gamma_{1i}X_{it} + \epsilon_{it}$

Level 2:  $\gamma_{0i} = \beta_{00} + \beta_{01}W_i + \nu_{0i}$

Level 2:  $\gamma_{1i} = \beta_{10} + \beta_{11}W_i + \nu_{1i}$

$W_i$  is the Level 2 variable which is normally distributed  $N(\mu_W^2, \sigma_W^2)$

AR(1) errors  $\epsilon_{it}$  with autocorrelation  $\rho$  and variance  $\sigma^2$

The distribution of the Level 1 variable:  $X_{it} = \mu_X + v_{0i} + \varepsilon_{it}$

$v_i$  is a Level 2 random effect which is normally distributed  $N(0, \sigma_{v_0}^2)$

AR(1) errors  $\varepsilon_{it}$  with autocorrelation  $\rho_\varepsilon$  and variance  $\sigma_\varepsilon^2$

Number of participants: introduce an increasing sequence of positive integers (comma-separated).

**Number of participants**

**Number of time points**

Figure 2: Ginette, can you please add the figure caption here?

## Step 2

**Set the value of the model parameters in the ApproxPowerIL application.**

Use the screenshots below to guide you through the process:

And, for the remainder of the parameters:

**Fixed intercept:  $\beta_{00}$**

42.98

**Effect of the Level 2 continuous variable  $W_i$  on the intercept:  $\beta_{01}$**

1.53

**Fixed slope:  $\beta_{10}$**

0.14

**Effect of the Level 2 continuous variable  $W_i$  on the slope:  $\beta_{11}$**

-0.01

**Standard deviation of Level 1 errors:  $\sigma$**

11.93

**Autocorrelation of Level 1 errors:  $\rho$**

0.43

**Standard deviation of random intercept:  $\sigma_{\nu_0}$**

12.86

**Standard deviation of random slope:  $\sigma_{\nu_1}$**

0.11

**Correlation between the random intercept and random slope:  $\rho_{\nu_{01}}$**

0.25

Figure 3: Ginette, can you please add the figure caption here?



Mean of Level 2 variable  $W_i$ :

Standard deviation of Level 2 variable  $W_i$ :

☒ Center the Level 2 variable  $W_i$

Mean of Level 1 variable  $X_{it}$ :

Standard deviation of the random intercept of the Level 1 variable  $X_{it}$ :

Standard deviation of Level 1 error of variable  $X_{it}$ :

Autocorrelation of Level 1 error of the variable  $X_{it}$ :

☒ Person-mean centered Level 1 variable  $X_{it}$  using the persons' mean

Select the tail of the hypothesis test:

Type I error:  $\alpha$

Figure 4: Ginette, can you please add the figure caption here?

### Step 3

#### Inspect the results.

Statistical power is higher than 90% when the number of participants is equal to or higher than 20.

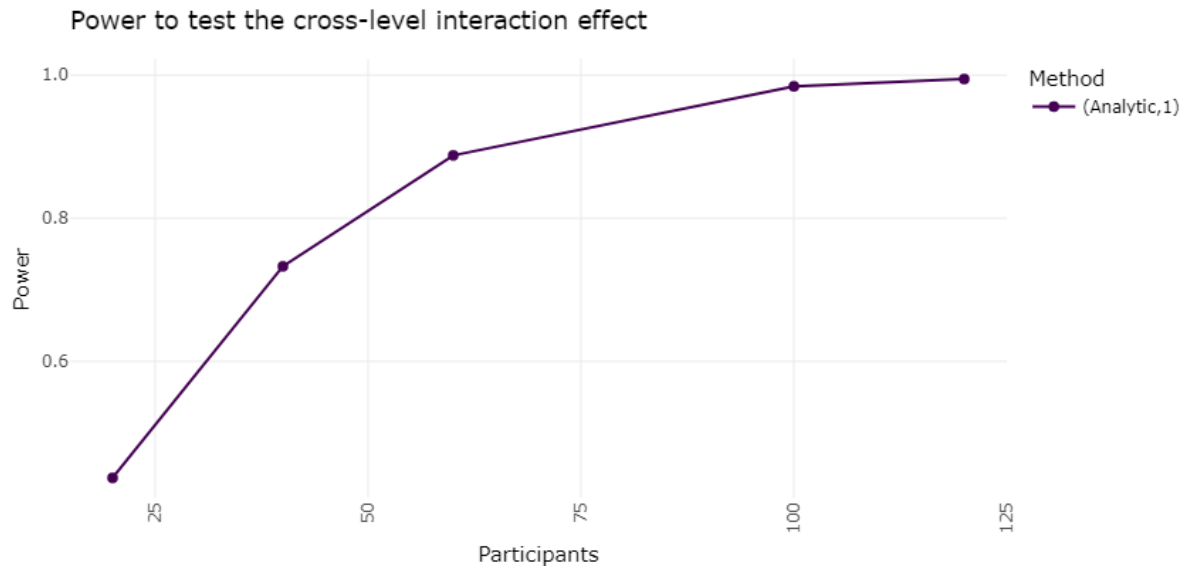


Figure 5: Ginette, can you please add the figure caption here?

### Simulation-based power analysis

To conduct the power analysis using the simulation-based approach we are going to use **PowerAnalysisIL**: a Shiny application and R package to perform power analysis to select the number of persons for multilevel models using the simulation-based approach.

download the app and run locally on their computer by executing the following The repository contains functions used in Lafit et al. (2021). Users can commands in R or RStudio at [PowerAnalysisIL](#).

```
# Install the app package from the `GitHub` repository.
library(devtools)
devtools::install_github("ginettelafit/PowerAnalysisIL", force = TRUE)

## Load package `PowerAnalysisIL`.
library(PowerAnalysisIL)
```

```
# Lunch the app from the `GitHub` gist.
shiny::runGist('6bac9d35c2521cc4fd91ce4b82490236')
```

## Step 1

Select the model and set the sample size in the PowerAnalysisIL application.

- Indicate the model of interest.
- Input the number of participants  $N$  (comma-separated):  $N = \{20, 40, 60, 80, 100\}$ .
- Input the number of repeated measurement occasions:  $T = 70$ .

**Choose a model (more information in panel About the Method):**

Model 7: Cross-level interaction effects (random slope) ▼

Model 7: Cross-level interaction effects (random slope)

Level 1:  $Y_{it} = \gamma_{0i} + \gamma_{1i}X_{it} + \epsilon_{it}$

Level 2:  $\gamma_{0i} = \beta_{00} + \beta_{01}W_i + \nu_{0i}$

Level 2:  $\gamma_{1i} = \beta_{10} + \beta_{11}W_i + \nu_{1i}$

$W_i$  is the level-2 variable which is normally distributed  $N(\mu_W^2, \sigma_W^2)$

AR(1) errors  $\epsilon_{it}$  with autocorrelation  $\rho_\epsilon$  and variance  $\sigma_\epsilon^2$

Number of participants: introduce an increasing sequence of positive integers (comma-separated).

**Number of participants**

20,40,60,100,120

**Number of time points**

70

Figure 6: Ginette, can you please add the figure caption here?

## Step 2

Set the value of the model parameters in the PowerAnalysisIL application.

Use the screenshots below to guide you through the process:

And, for the remainder of the parameters:

**Fixed intercept:  $\beta_{00}$**

42.98

**Effect of the level-2 continuous variable on the intercept:  $\beta_{01}$**

1.53

**Fixed slope:  $\beta_{10}$**

0.14

**Effect of the level-2 continuous variable on the slope:  $\beta_{11}$**

-0.01

**Standard deviation of level-1 errors:  $\sigma_{\epsilon}$**

11.93

**Autocorrelation of level-1 errors:  $\rho_{\epsilon}$**

0.43

**Standard deviation of random intercept:  $\sigma_{\nu_0}$**

12.86

**Standard deviation of random slope:  $\sigma_{\nu_1}$**

0.11

**Correlation between the random intercept and random slope:**  
 $\rho_{\nu_{01}}$

0.25

Figure 7: Ginette, can you please add the figure caption here?

Mean of time-varying variable X:

Standard deviation of time-varying variable X:

☒ Person mean centering  $X_{it}$  using the individual mean

Mean of level-2 variable W:

Standard deviation of level-2 variable W:

☒ Center the level-2 variable W

☒ Estimate AR(1) correlated errors  $\epsilon_{it}$

Type I error:  $\alpha$

Monte Carlo Replicates

Choose the method to fit linear mixed-effects model

Figure 8: Ginette, can you please add the figure caption here?

### Step 3

#### Inspect the results.

Statistical power is higher than 90% when the number of participants is equal to or higher than 20.

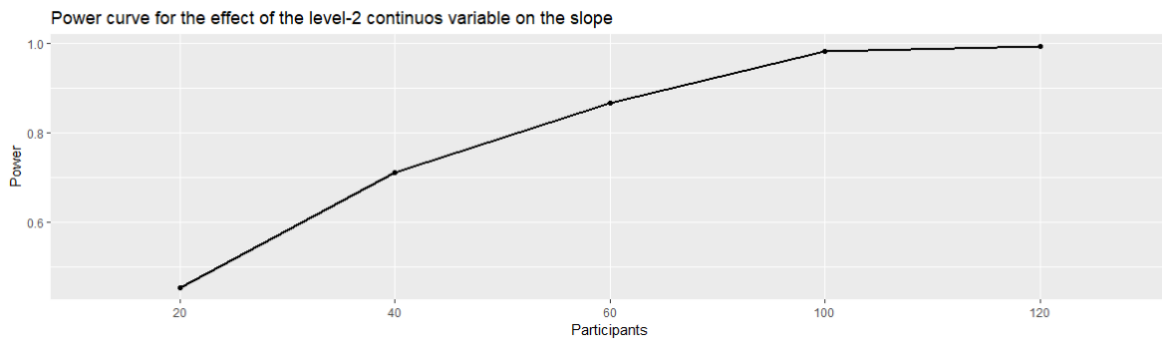


Figure 9: Ginette, can you please add the figure caption here?

### Session information

Using the command below, we can print the `session` information (i.e., operating system, details about the R installation, and so on) for reproducibility purposes.

```
# Session information.  
sessionInfo()
```

```
R version 4.3.0 (2023-04-21)  
Platform: aarch64-apple-darwin20 (64-bit)  
Running under: macOS Ventura 13.4
```

```
Matrix products: default  
BLAS: /Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/lib/libRblas.0.dylib  
LAPACK: /Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/lib/libRlapack.dylib;
```

```
locale:  
[1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
```

```
time zone: Europe/Amsterdam  
tzcode source: internal
```

attached base packages:

```
[1] stats      graphics  grDevices  utils      datasets  methods   base
```

loaded via a namespace (and not attached):

```
[1] compiler_4.3.0 fastmap_1.1.1 cli_3.6.1      tools_4.3.0  
[5] htmltools_0.5.5 rstudioapi_0.14 yaml_2.3.7      rmarkdown_2.22  
[9] knitr_1.43      jsonlite_1.8.5 xfun_0.39       digest_0.6.31  
[13] rlang_1.1.1     evaluate_0.21
```

## References

- Goldstein, H., Healy, M. J., & Rasbash, J. (1994). Multilevel time series models with applications to repeated measures data. *Statistics in Medicine*, 13(16), 1643–1655.
- Hamilton, J. D. (1994). *Time series analysis* (Vol. 2). Princeton New Jersey.
- Lafit, G., Adolf, J. K., Dejonckheere, E., Myin-Germeys, I., Viechtbauer, W., & Ceulemans, E. (2021). Selection of the number of participants in intensive longitudinal studies: A user-friendly shiny app and tutorial for performing power analysis in multilevel regression models that account for temporal dependencies. *Advances in Methods and Practices in Psychological Science*, 4(1), 2515245920978738.
- Lafit, G., Artner, R., & Ceulemans, E. (2023). *Enabling analytical power calculations for multi-level models with autocorrelated errors through deriving and approximating the information matrix*.