

gTheoryShiny: An online application for interactive G-theory inference

Jihong Zhang

Introduction

gTheoryShiny is an *R/Shiny* application for interactive construction, and inference of generalizability theory (also named G-theory) for performance assessment, which provide friendly GUI for users lacking of programming skills. It's heavily based on **lme4** (Bates et al., 2015) for multilevel modeling. To illustrate the functionality of this app, this manual is organized as following:

- Input section
 1. Individual Data (.csv)
- Facet screening
- Data Analysis
 1. g-study estimation and bootstrapping error
 2. d-study estimation and bootstrapping error

Input section

1. Individual Data (.csv)

Currently, the app also support .csv file. The data supported by gTheoryShiny can be either long-format or wide-format.

Wide-format data

If your data is wide-format, the app can help you transform the wide-format data into the long-format data. First, please make sure the first several rows of data include identifier of all facets (in the example table below, facets are *task1/task2* and *rater1/rater2*). Header is optional. If the data contains the header, check the box of *include header*. The first column of the data should be subject ID (*Person* in the table). Then, each row should represent each subject's all observations.

In addition, there are three settings you should fill in: (1) how many rows represent tag/ID, (2) the pre-fix of tag/ID, (3) column names of tag/ID in the transformed long-format data. For example, to transform table 1 to table 2, setting 1 is 2 (first two rows contains all facets' identifier), setting 2 can be *T;R* (the identifiers of facets have pre-fix T and R with the same order with the order of each facet's rows), and setting 3 can be *Task;Rater* (in long-format, the column names of facets are Task and Rater in order).

It should also be noted that the long-format data transformed in this way automate generate two columns in the first column and the last column respectively: *ID* and *Score*. ID's values correspond to subject ID, and Score's values are same as values in the corresponding cells in wide-format data. Currently, there's no way to change the column names of these two variables.

Finally, don't forget to click *Transform* button to finish the wide-to-long transformation.

Table 1: An Example of Wide-format Data

Person	Task1.Rater1	Task1.Rater2	Task2.Rater1	Col5
	Task1	Task1	Task2	Task2
	Rater1	Rater2	Rater1	Rater2
1	5	6	7	8

Table 2: An Example of Long-format Data

ID	Task	Rater	Score
1	T_Task1	R_Rater1	5
1	T_Task1	R_Rater2	6
1	T_Task2	R_Rater1	7
1	T_Task2	R_Rater2	8

Long-format Data

If the data is already long-format before uploading, there are no extra step to set up. As figure 1 shows, only things needed to be checked are (1) check *long-format* check box under the Data

property, (2) check or un-check *include header* depending on whether the data has headers or not.

Choose CSV File:

Upload... Brennan.3.2Long.cs

Upload complete

Notice: If data is wide-format, make sure the first two rows of your data file should be TAG/ID, first column should be subject ID

Data property:

long-format

include header

Raw Data:

Show 10 entries Search:

	Task ↕	Person ↕	Rater ↕	Score ↕
1	1	1	1	5
2	1	2	1	9
3	1	3	1	3
4	1	4	1	7
5	1	5	1	9
6	1	6	1	3
7	1	7	1	7
8	1	8	1	5
9	1	9	1	9
10	1	10	1	4

Showing 1 to 10 of 120 entries

Previous 1 2 3 4 5 ... 12 Next

Figure 1: Long-format data input

Facet Screening

The facet screening is second step for g-theory data analysis. When data is long-format, users should specify which column corresponds to (1) ID, (2) facets, or (3) outcome. Then two tables will pop up: (1) Structural table, which represents the relations between facets (nested or crossed), and (2) Summary table, which includes the sample size for each condition (all combination of facets).

Figure 2 shows an example of facet screening setup. Select *Person* as ID, then check *Task* and *Rater* as facets, finally select *Score* as outcome. In Structural table, it shows that *Task* and *Rater* are nested in this example.

Control facets/outcome:

Which column represents ID:

Person

Which column(s) represent facets:

Task

Person

Rater

Score

Which column represents outcome:

Score

confirm

Structural Table:

Show 10 entries Search:

	f1	f2	NestedOrCrossed
1	Task	Rater	Nested

Showing 1 to 1 of 1 entries Previous 1 Next

Summary Table:

Show 10 entries Search:

	Task	Rater	Sample Size (Outcome)
1	1	1	5
2	1	2	7
3	1	3	6
4	1	4	4
5	2	5	4
6	2	6	5
7	2	7	4
8	2	8	4
9	3	9	6
10	3	10	5

Showing 1 to 10 of 12 entries Previous 1 2 Next

Figure 2: Facet screening setup

Data Analysis

After users finish facet screening, recommended formula is automate generated in Data Analysis tab page. Figure 3 shows an example of g-thoery estimation setup. The formula follows multilevel modeling formula. The left hand side of equal sign represents the outcome. The right hand side contains the linear combination of facets and ID. For example,

$$Y = PersonID + Facet1 : Facet2 + Facet2 + Facet3$$

where Facet1 is nested within Facet2, Facet2 and Facet 3 are crossed, and PersonID denotes subject ID.

Users are allowed to specify their own formula for g-theory estimation following the same format as above.

Link Function option allows users to specify the relation between outcome and linear combination of facets. Possible choices include identity, logit, probit, poisson etc. Currently, only identity is supported.

Number of bootstrap allows users to specify how many bootstrapping iterations for bootstrapping measurement error estimation. The default is 200.

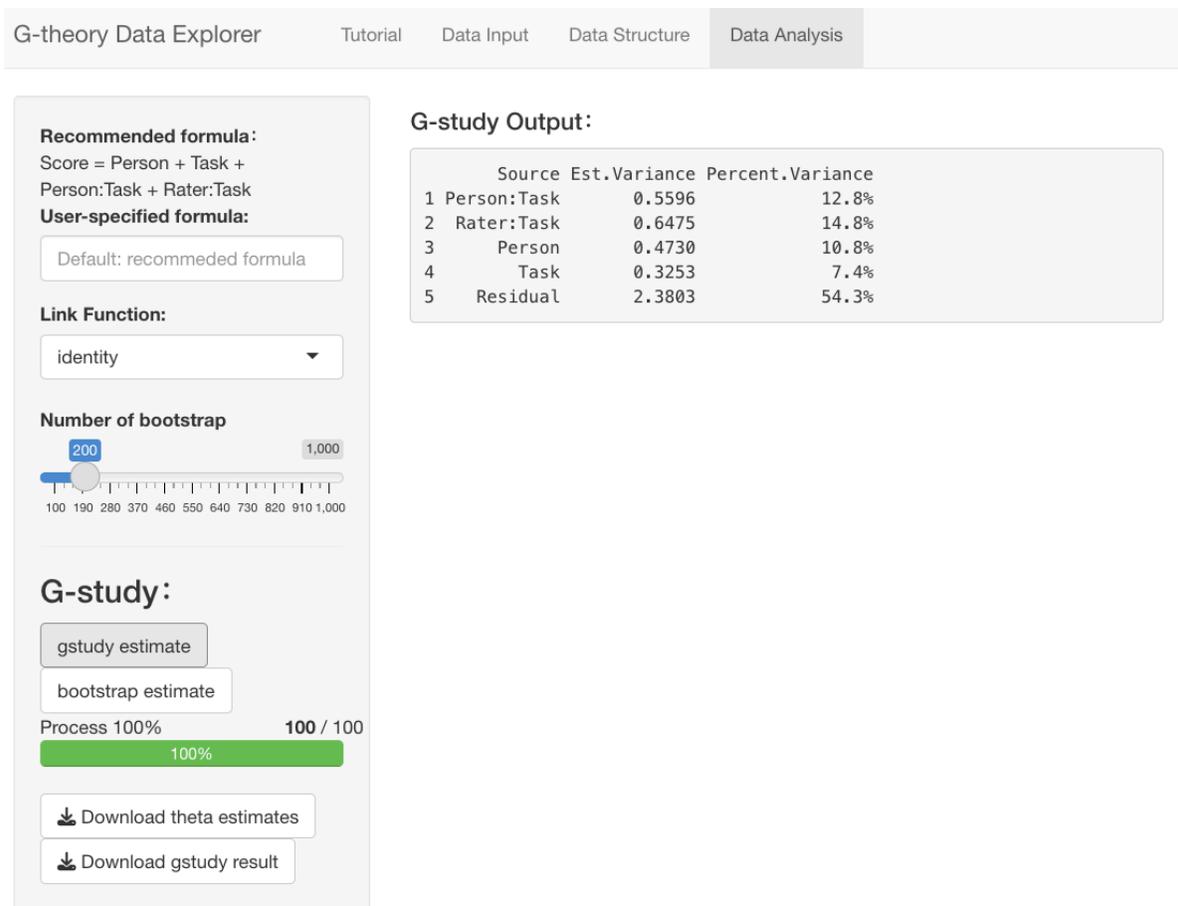


Figure 3: Screenshot of g-theory estimation

1. G-theory estimation and bootstrapping

Clicking *gstudy estimate* button will provide the g-study results including sources of variance components, the estimates of variance, and percent of each variance component.

Then, as shown in Figure 4, users can click *bootstrap estimate* button to generate bootstrapping standard deviation of g-study results. All results of g-study are downloadable.

The screenshot displays the G-study interface with the following components:

- Recommended formula:** Score = Person + Task + Person:Task + Rater:Task
- User-specified formula:** Default: recommended formula
- Link Function:** identity
- Number of bootstrap:** A slider set to 200, ranging from 100 to 1,000.
- G-study:** Includes buttons for 'gstudy estimate', 'bootstrap estimate', and a progress bar showing 'Process 100%'.
- Download buttons:** 'Download theta estimates', 'Download gstudy result', and 'Download bootstrap result'.
- G-study Output Table:**

	Source	Est.Variance	Percent.Variance
1	Person:Task	0.5596	12.8%
2	Rater:Task	0.6475	14.8%
3	Person	0.4730	10.8%
4	Task	0.3253	7.4%
5	Residual	2.3803	54.3%
- Estimate Bootstrapping SD for G-study Table:**

	Source	Est.Variance	Percent.Variance	2.5%	97.5%
1	Person:Task	0.5596	12.8%	0.0000000	1.350828
2	Rater:Task	0.6475	14.8%	0.0162325	1.534110
3	Person	0.4730	10.8%	0.0000000	1.337575
4	Task	0.3253	7.4%	0.0000000	2.354895
5	Residual	2.3803	54.3%	1.7227850	3.071122

Figure 4: Screenshot of bootstrap estimation

2. D-study estimation and bootstrapping

To get results of D-study, users need to select (1) which facet to change, (2) how many potential levels of facets. Users can change other facet by repeating these two steps, then click *confirm facet levels* button. For example, the figure below shows facet *Task* is set up to 60 level in D-study.

After settings of d-study are finished, clicking *dstudy estimate* will provide the results of dstudy, which contains two parts: levels of facets in D-study and statistics of d-study. Figure 6 shows an example of d-study with there are 60 tasks in d-study.

Similar to G-study, to get bootstrapping SD of d-study, click *bootstrap estimate* button. Finally, click *download* button to download results of D-study.

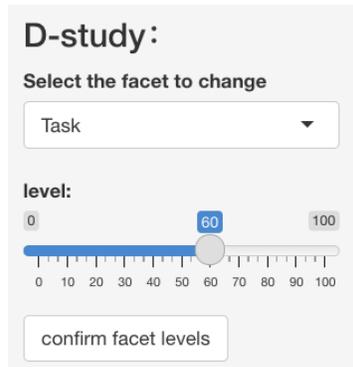


Figure 5: Screenshot of D-study settings

D-study Output:

Sample Size:

Task	Rater
60	12

Result:

The generalizability coefficient is: 0.9739873.
 The dependability coefficient is: 0.9614728.
 The relative error is: 0.01263264.
 The absolute error is: 0.01895361.

.

	Source	Est.Variance	N	Est.(Var/N)
1	Person:Task	0.5596	60	0.0093266667
2	Rater:Task	0.6475	720	0.0008993056
3	Person	0.4730	120	0.4730000000
4	Task	0.3253	60	0.0054216667
5	Residual	2.3803	720	0.0033059722

Figure 6: Screenshot of D-study output