

# Package ‘resquin’

September 11, 2024

**Title** Response Quality Indicators for Survey Research

**Version** 0.0.2

**Description** Calculate common survey data quality indicators for multi-item scales and matrix questions. Currently supports the calculation of response style indicators and response distribution indicators. For an overview on response quality indicators see Bhaktha N, Henning S, Clemens L (2024). 'Characterizing response quality in surveys with multi-item scales: A unified framework' <<https://osf.io/9gs67/>>.

**License** GPL (>= 3)

**Encoding** UTF-8

**RoxygenNote** 7.3.2

**Imports** cli, purrr, vctrs

**Suggests** knitr, rmarkdown, testthat (>= 3.0.0)

**Config/testthat/edition** 3

**VignetteBuilder** knitr

**URL** <https://github.com/MatRoth/resquin>,  
<https://matroth.github.io/resquin/>

**BugReports** <https://github.com/MatRoth/resquin/issues>

**NeedsCompilation** no

**Author** Matthias Roth [aut, cre, cph] (<<https://orcid.org/0000-0003-4369-8106>>),  
Nivedita Bhaktha [aut, ctb],  
Matthias Bluemke [aut, ctb],  
Thomas Knopf [aut, ctb],  
Fabienne Krämer [aut, ctb],  
Clemens Lechner [aut, ctb],  
Çağla Yildiz [aut, ctb]

**Maintainer** Matthias Roth <[matthias.roth@gesis.org](mailto:matthias.roth@gesis.org)>

**Repository** CRAN

**Date/Publication** 2024-09-11 15:10:02 UTC

## Contents

resp_distributions . . . . .	2
resp_styles . . . . .	4

<b>Index</b>	<b>7</b>
--------------	----------

---

resp_distributions	<i>Compute response distribution indicators</i>
--------------------	---

---

### Description

Compute response distribution indicators for responses to multi-item scales or matrix questions.

### Usage

```
resp_distributions(x, min_valid_responses = 1)
```

### Arguments

x	A data frame containing survey responses in wide format. For more information see section "Data requirements" below.
min_valid_responses	numeric between 0 and 1. Defines the share of valid responses a respondent must have to calculate response quality indicators. Default is 1.

### Details

The following response distribution indicators are calculated per respondent:

- n\_na: number of intra-individual missing answers
- prop\_na: proportion of intra-individual missing responses
- ii\_mean: intra-individual mean
- ii\_median: intra-individual median
- ii\_sd: intra-individual standard deviation
- mahal: mahalanobis distance per respondent.

Intra-individual response variability (ii\_sd) has been proposed to measure insufficient effort responding (Dunn et al., 2018) and to distinguish between random and conscientious responding (Marjanovic et al, 2015).

Intra-individual location indicators can be used to assess the average location of responses on a set of questions (ii\_mean, ii\_median).

Mahalanobis distance is an outlier detection indicator. It represents the distance of a participant's responses from the center of a multivariate normal distribution defined by the data of all respondents.

**Value**

Returns a data frame with response quality indicators per respondent. Dimensions:

- Rows: Equal to number of rows in *x*.
- Columns: Six, one for each response distribution indicator.

**Data requirements**

`resp_distributions()` assumes that data comes from multi-item scales or matrix questions, which have the same number and labeling of response options for many questions. The input data frame must be structured in the following way:

- The data frame is in wide format, meaning each row represents one respondent, each column represents one variable.
- All responses have integer values.
- Missing values are set to NA.

**Reverse coding of variables**

The interpretation of the indicators depends on the whether response data of negatively worded questions was reversed or not:

- Do not reverse data of negatively worded questions if you want to assess average response patterns (Dunn et al., 2018).
- Reverse data of negatively worded questions if you want to assess whether responses are distributed randomly or not with respect to an assumed latent variable (Marjanovic et al., 2015).

**Mahalanobis distance could not be calculated**

Under certain circumstances, the mahalanobis distance can not be calculated. This may be if there is high collinearity (correlation between variables) or if there are too many missing values. Although this can happen in survey research data, this message can also indicate that something in the data is "off" due to one of the reasons stated above. A manual inspection for low-quality responses can be a next step.

**Author(s)**

Matthias Roth, Matthias Bluemke & Clemens Lechner

**References**

Dunn, Alexandra M., Eric D. Heggstad, Linda R. Shanock, and Nels Theilgard. 2018. "Intra-Individual Response Variability as an Indicator of Insufficient Effort Responding: Comparison to Other Indicators and Relationships with Individual Differences." *Journal of Business and Psychology* 33(1):105–21. doi: 10.1007/s10869-016-9479-0.

Marjanovic, Zdravko, Ronald Holden, Ward Struthers, Robert Cribbie, and Esther Greenglass. 2015. "The Inter-Item Standard Deviation (ISD): An Index That Discriminates between Conscientious and Random Responders." *Personality and Individual Differences* 84:79–83. doi: 10.1016/j.paid.2014.08.021.

**See Also**

[resp\\_styles\(\)](#) for calculating response style indicators.

**Examples**

```
# A small test data set with ten respondents
# and responses to three survey questions
# with response scales from 1 to 5.
testdata <- data.frame(
  var_a = c(1,4,3,5,3,2,3,1,3,NA),
  var_b = c(2,5,2,3,4,1,NA,2,NA,NA),
  var_c = c(1,2,3,NA,3,4,4,5,NA,NA))

# Calculate response distribution indicators
resp_distributions(x = testdata) |>
  round(2)

# Include respondents with NA values by decreasing the
# necessary number of valid responses per respondent.

resp_distributions(
  x = testdata,
  min_valid_responses = 0.2) |>
  round(2)
```

---

resp\_styles

*Compute response style indicators*

---

**Description**

Calculates response style indicators for matrix questions or multi-item scales.

**Usage**

```
resp_styles(x, scale_min, scale_max, min_valid_responses = 1, normalize = TRUE)
```

**Arguments**

x	A data frame containing survey responses in wide format. For more information see section "Data requirements" below.
scale_min	numeric. Minimum of scale provided.
scale_max	numeric. Maximum of scale provided.
min_valid_responses	numeric between 0 and 1. Defines the share of valid responses a respondent must have to calculate response style indicators.
normalize	logical. If <i>TRUE</i> , counts of response style indicators will be divided by the number of non-missing responses per respondent. Default is <i>TRUE</i> .

## Details

Response styles capture systematic shifts in respondents response behavior. `resp_styles()` is aimed at multi-item scales or matrix questions which use the same number of response options for many questions.

#' The following response style indicators are calculated per respondent: Middle response style (MRS), acquiescence response style (ARS), disacquiescence response style (DARS), extreme response style (ERS) and non-extreme response style (NERS).

The response style indicators are calculated in the following way

- MRS: Sum of mid point responses.
- ARS: Sum of responses larger than midpoint.
- DARS: Sum of responses lower than midpoint.
- ERS: Sum of lowest or highest category responses.
- NERS: Sum of responses between lowest and highest response category.

Note that ARS and DRS assume that the polarity of the scale is positive. This means that higher numerical values indicate agreement and lower numerical values indicate disagreement. MRS can only be calculated if the scale has a numeric midpoint.

Also note that the response style literature is fragmented (Bhaktha et al., 2024). Response styles calculated with `resp_styles()` are based on van Vaerenbergh & Thomas (2024). However, we used the name non-extreme response style (NERS) instead of mild response style, to emphasize that NERS is the inverse of ERS. Both appear in the literature (for a NERS example see Wetzel et al. (2013)). Consult literature in your field of research to find appropriate names for the response style indicators calculated here.

## Value

Returns a data frame with response style indicators per respondent.

- Rows: Equal to number of rows in `x`.
- Columns: Five, one for each response style indicator.

## Data requirements

`resp_styles()` assumes that the input data frame is structured in the following way:

- The data frame is in wide format, meaning each row represents one respondent, each column represents one variable.
- The variables are in same the order as the questions respondents saw while taking the survey.
- Reverse keyed variables are in their original form. No items were recoded.
- All responses have integer values.
- Questions have the same number of response options.
- Missing values are set to NA.

## Author(s)

Matthias Roth, Matthias Bluemke & Clemens Lechner

## References

Bhaktha, Nivedita, Henning Silber, and Clemens Lechner. 2024. „Characterizing response quality in surveys with multi-item scales: A unified framework“. OSF-preprint: <https://osf.io/9gs67/> van Vaerenbergh, Y., and T. D. Thomas. 2013. „Response Styles in Survey Research: A Literature Review of Antecedents, Consequences, and Remedies“. *International Journal of Public Opinion Research* 25(2):195–217. doi: 10.1093/ijpor/eds021. Wetzels, Eunike, Claus H. Carstensen, and Jan R. Böhnke. 2013. „Consistency of Extreme Response Style and Non-Extreme Response Style across Traits“. *Journal of Research in Personality* 47(2):178–89. doi: 10.1016/j.jrp.2012.10.010.

## See Also

[resp\\_distributions\(\)](#) for calculating response distribution indicators.

## Examples

```
# A test data set with ten respondents
# and responses to three survey questions
# with response scales from 1 to 5.
testdata <- data.frame(
  var_a = c(1,4,3,5,3,2,3,1,3,NA),
  var_b = c(2,5,2,3,4,1,NA,2,NA,NA),
  var_c = c(1,2,3,NA,3,4,4,5,NA,NA))

# Calculate response distribution indicators
resp_styles(testdata,
            scale_min = 1,
            scale_max = 5) |>
  round(2) # round to second decimal

# Include respondents with NA values by decreasing the
# necessary number of valid responses per respondent.
resp_styles(testdata,
            scale_min = 1,
            scale_max = 5,
            min_valid_responses = 0.2) |>
  round(2) # round to second decimal

# Get counts of responses attributable to response styles.
resp_styles(testdata,
            scale_min = 1,
            scale_max = 5,
            normalize = FALSE)
```

# Index

`resp_distributions`, 2  
`resp_distributions()`, 6  
`resp_styles`, 4  
`resp_styles()`, 4